Past & Contemporary Patterns of Marine Biodiversity: Making Inferences from Anadromous Sea Lamprey Population Genomics & Publicly Archived Genetic Data

by

Eleana Karachaliou

A Thesis submitted to the Faculty of Graduate Studies of

The University of Manitoba

in partial fulfilment of the requirements of the degree of

DOCTOR OF PHILOSOPHY

Department of Biological Sciences

University of Manitoba

Winnipeg

ABSTRACT

Maintaining healthy oceans and the services they provide relies on safeguarding marine biodiversity. Genetic diversity is the foundation of biodiversity, underlying adaptive potential and population persistence, ecosystem resilience and stability. Planning for future genetic diversity conservation will be multifaceted. My thesis aims to uncover both patterns and drivers of global genetic diversity in marine fishes and the population genetics of a single species, taking a simultaneously global and single species approach to conservation genetics in the marine environment. I initially used a synthesis methodology, taking advantage of the large amount of publicly available genetic data. My aim was to determine global patterns of genetic diversity for marine fishes and their drivers. I found that global patterns of genetic diversity seemingly follow a west to east direction along global coastlines. I also uncovered human-caused declines in genetic diversity but not population differentiation. I detected variation in genetic diversity and gene flow levels connected to oceanographic parameters. I also found that species richness is correlated with higher genetic diversity, more genetically similar populations and lower effective population size. I then focused on a single species, anadromous sea lamprey *Petromyzon marinus.* The aim of my study was to apply a conservation genetics approach to this widely distributed species. As part of my study, I uncovered sea lamprey population genetic structure, status, demographic and evolutionary history, and likelihood of future maladaptation. I identified distinct sea lamprey management units across the species' native range. I identified locations at which sea lamprey are vulnerable, both currently and under future climate change scenarios. I reconstructed deep demographic history and timed the split between sea lamprey populations on the two Atlantic coasts. I also uncovered that in recent evolutionary time, sea lamprey have faced population declines connected to human activity. My research has extended understanding of

genetic diversity across global marine fish species, sea lamprey genetic structure, evolutionary trajectory and vulnerability across their native range. I interpret my results from a conservation genetics point of view, aiming to inform efforts for biodiversity conservation and show the importance in accounting for both the detail and the greater picture.

Table of Contents

LIST OF TABLES

LIST OF FIGURES

CONTRIBUTIONS OF AUTHORS

Chapter 2: Conceptualization - E.K., C.S., M.F.D., C.J.G.; Methodology - E.K., C.S., E.d.G., C.M., C.J.G.; Analysis - E.K.; Writing the First Draft – E.K.; Revision and Editing - E.K., M.F.D., C.J.G.

Chapter 3: Conceptualization - E.K., C.J.G., M.F.D.; Methodology - E.K., J.S., E.d.G., C.J.G.; Analysis - E.K.; Writing the First Draft – E.K.; Revision and Editing - E.K., J.S., C.J.G., M.F.D.

Chapter 4: Conceptualization - E.K., C.J.G., M.F.D.; Methodology - E.K., J.S., E.d.G., C.J.G.; Analysis - E.K.; Writing the First Draft – E.K.; Revision and Editing - E.K., J.S., C.J.G., M.F.D.

ACKNOWLEDGEMENTS

Many thanks to my supervisors Margaret Docker and Colin Garroway for their support and encouragement during my PhD journey. I would also like to thank my committee members Olivier Tremblay-Savard and Stephen Petersen for their guidance. I owe many thanks to my family and friends for their moral support and for always believing in me. Many thanks also to Athanasios Exadactylos, my professor during my BSc degree, who encouraged me to pursue science.

Open data practices were an important part of my project, and I am grateful to the authors who made their data publicly available. I would also like to thank my collaborators who provided me with anadromous sea lamprey samples - Ted Castro-Santos, Mike Wilkie, Guillaume Evanno, Magnús Jóhannsson, Benóný Jónsson, Jan Baer, Fiona Bracken, Catarina Mateus, Thomas Evans, Michael Fisk, Jeremy McCargo, Gabriela M. Hogue, Mikael Svensson and Elisabeth Thysell. Thanks also to Chloé Schmidt, Jaanus Suurväli and Evelien de Greef for the technical support they provided as part of this project.

This work was supported by NSERC Discovery Grants awarded to C.J.G and M.F.D., a University of Manitoba Faculty of Science Research Chair award to M.F.D., and a Faculty of Science Enhancement of Grant Stipends award to E.K.

GLOSSARY

admixture – interbreeding between two previously isolated populations

gene flow – movement of genetic material between populations

genetic diversity – variation in genetic composition within a population or species

gene diversity (expected heterozygosity) - probability that an individual will be heterozygous at a given locus

allelic richness – number of alleles observed in a population, effectively an allele count

species richness – number of species found within a distinct region

FST - between-population genetic differentiation

effective population size (Ne) - rate at which a population is susceptible to genetic drift

genetic offsets - metric of genome-environment mismatch under future climate change

allelic turnover - turnover of allele frequencies across environmental gradients

CHAPTER 1: INTRODUCTION

From sustenance to culture, our lives are interconnected with the world's oceans. An increasingly large percentage of the human diet worldwide depends on marine resources (Costello *et al.* 2020; FAO 2022; Jouffray *et al.* 2020), with developing countries disproportionately dependent on fisheries compared to developed countries (FAO 2022). Natural processes present across global oceans can also help mitigate climate change. Marine species such as mangroves, seagrasses and coral reefs can buffer against the effects of climate change, helping prevent coastal erosion and protecting coastal communities from extreme weather events (Gracia *et al.* 2018; Guannel *et al.* 2016). The biological carbon pump, a process where the ocean functions as a carbon sink absorbing carbon that would otherwise contribute to climate change, relies on biological processes of marine organisms (Heinze *et al.* 2015). Marine species also hold the possibility of discovering new materials that could contribute to global sustainability goals (Maeda *et al.* 2018; Tennakoon *et al.* 2023). Therefore, sustainability across global oceans will not be possible without protecting the organisms that inhabit them. Preserving marine biodiversity by enabling the perseverance of complex marine ecosystems and their dynamics should be a priority when planning for a sustainable future and rethinking our relationship with nature. Despite the culture and values of many coastal communities being deeply connected to the ocean, as human societies become increasingly detached from nature, our relationship with the ocean is often disregarded. The benefits and ethics of maintaining healthy oceans mean that the value of undoing centuries of human-caused degradation of the marine environment should, in theory, be indisputable.

Protecting marine biodiversity means maintaining genetic diversity at sustainable levels and facilitating species' resilience by ensuring their ability to adapt (Barrett & Schulter 2008).

Meanwhile, contemporary genetic diversity is estimated to have declined by around 6% across global taxa, including marine fishes (Leigh *et al.* 2019). Overfishing has caused genetic diversity declines in commercially harvested marine fishes (Pinsky & Palumbi 2014). Understanding the extent of human impacts and the dynamics of underlying natural processes across global oceans is needed to halt known declines in both biodiversity (Barnosky *et al*. 2011; Sala & Knowlton 2006) and genetic diversity (Leigh *et al*. 2019; Pinsky & Palumbi 2014). As the results of human-caused environmental changes become increasingly noticeable, conservation of genetic diversity should be at the forefront of planning for a sustainable future. Included in the Kunming-Montreal Global Biodiversity Framework's goals for 2050 is safeguarding genetic diversity and subsequent adaptive potential across populations of wild and domesticated species. Alongside the necessity of conservation actions, the framework highlights the need for equitable sharing of genetic resources. Equitable sharing of these resources and the rights and benefits that derive from them is a step forward in ensuring positive outcomes are felt globally. This way, involvement will be encouraged across countries and communities, which is important in encouraging global action to prevent biodiversity declines and working towards achieving genetic diversity conservation goals. Efforts to maintain genetic diversity are relatively new, and they provide an opportunity to ensure conservation goals and planning are truly science-based, equitable and global.

Producing high-quality scientific research to support and inform policy is a baseline for meeting conservation goals. There is a need for increased inclusion of scientists in the decision-making process and a true intersection of science and policy. Plans underway to create a global ocean observation framework based on metrics applicable to marine policy and management support the inclusion of genetic diversity as a measure of biodiversity (Muller-Karger *et al.*

2018). Restoration and maintenance of genetic diversity is explicitly included in the Kunming-Montreal Global Biodiversity Framework for 2030. The proportion of populations within a species with an effective population size, the rate at which a population is susceptible to genetic drift, >500 is listed as a headline indicator. However, there is room for improvement, with more details needed on how genetic diversity will be restored and monitored, data limitations addressed, and the use of genetic diversity metrics as indicators (Hoban *et al.* 2020; Hughes & Grumbine 2023). Genetic diversity metrics are equally valuable in a system-specific context. An example of applying genetic tools to system-specific management needs is the long-term genetic monitoring programs for managing ecologically and culturally important diadromous species in the Columbia River Basin (USA) (Hand *et al.* 2018; Hess *et al.* 2022). Optimizing and applying genetic diversity metrics for integration into policy has been extensively reviewed, with global implementation still underway but substantial efforts at the country level (Hoban *et al.* 2022). As awareness of the importance of genetic diversity monitoring and reporting increases, so does the need for good data practices. Ensuring genetic data is user accessible, reproducible and applicable across studies will encourage its widespread use (Leigh *et al.* 2023).

Conservation needs across marine fishes can be hard to define, with different strategies required at the species, population and stock level for commercially harvested species, to account for different threat levels and natural processes experienced across their range. Defining biological populations and management units can be challenging in the marine environment where there is frequent mismatch between the two (Reiss *et al.* 2009), and population genetic studies are a valuable resource for addressing this problem (Hemmer-Hansen *et al.* 2019; Spies & Punt 2015). Population vulnerability assessment using effective population size (Ne) as a metric is also a valuable conservation tool in the marine environment (Hare *et al.* 2011). One of

the most commonly used databases for informing conservation priorities is the International Union for the Conservation of Nature (IUCN) Red List which estimates species extinction risk (IUCN 2012). However, the IUCN Red List does not directly include genetic diversity despite its widespread use in decision-making. Despite opposing suggestions, it cannot currently be used to address issues regarding conservation of genetic diversity, as genetic diversity levels do not reflect IUCN threat status on a species-by-species basis (Schmidt *et al.* 2023). Producing genetic data on population structure and vulnerability that is directly applicable to conservation and management will help address potential mismatch between biodiversity conservation strategies.

Subject to generations of overfishing, pollutants and the effects of urbanization since humans started organizing as societies, evidence of human intervention in the marine environment is expected to be easily detected. Simultaneously, marine fishes display a large range of life histories and inhabit oceanographically diverse environments. Despite spatial variation in human-caused impact levels, diverse life histories and habitats, past climatic processes and contemporary declines in biodiversity due to human activity are evident at the genetic level (Buckley *et al.* 2022; de Lafontaine *et al.* 2018; Karachaliou *et al.* 2025; Layton *et al.* 2021; Lehnert *et al.* 2019; Pinsky & Palumbi 2014). Therefore, marine fishes can provide valuable insights into both human-caused and environment-driven evolution on ecological time scales (Karachaliou *et al.* 2025; Layton *et al.* 2021; Lehnert *et al.* 2019; Pinsky & Palumbi 2014). Determining the genetic basis of conditions affecting marine life across species and populations could enable the isolation of natural evolutionary mechanisms from those caused by human interference and climate change. Global biodiversity conservation and management design needs to not just focus on the species level but also consider the unique needs and characteristics of each population, overall structure and inter-population dynamics. Distinct

populations within fish species can contribute towards resilience and persistence under unstable environmental conditions due to per-population life history and local adaptation (Hilborn *et al.* 2003). Identifying populations' current status and predicting future responses under varying oceanographic conditions linked to climate change could assist with understanding future trends and planning for future safeguarding of genetic diversity.

Increased accessibility of next-generation sequencing techniques provides an opportunity to better understand evolutionary processes underlying oceanographically complex marine environments (e.g., Barceló *et al.* 2022; Burt *et al.* 2024; Van Wyngaarden *et al.* 2018). Advances in genetics have made a wide array of methods and resources available to deepen understanding of our study systems. We can now readily search for genetic variants across the genome in non-model organisms (Andrews *et al.* 2016; Ellegren 2014). Past schools of thought claimed that marine fish species are not easily susceptible to genetic change due to their wide distribution, large populations, high dispersal and connectivity, but advances in genetics are challenging these beliefs (Hauser & Carvalho 2008). Contemporary studies are proving that marine systems are more complex than previously assumed, uncovering finer-scale neutral and adaptive population structure with implications for fisheries management (e.g., Canales-Aguirre *et al.* 2022; Limborg *et al.* 2012; Longo *et al.* 2020; Milano *et al.* 2014). This approach also holds in lampreys and other culturally and economically important diadromous species where there are signals of population structure (e.g., Bourret *et al.* 2013; Candy *et al.* 2015; Clemens *et al.* 2017; Hess *et al.* 2020). Population genetics approaches can benefit management and conservation planning in the marine environment. Examples include designing marine protected areas (Xuereb *et al.* 2019), genetic rescue (Neely *et al.* 2021), population status assessment (Hare *et al.* 2011), identifying management units (Dahle *et al.* 2018; Palsbøll *et al.* 2007) and

determining harvesting quotas (Dahle *et al.* 2018). When possible, next generation sequencing data should be preferred over more limited genetic markers, its utility already demonstrated in lamprey genetic monitoring and management (Hess *et al.* 2013; 2020; 2022). Meanwhile, Single Nucleotide Polymorphism (SNP) panels have also been used to uncover demography, selection and evolutionary history in other anadromous species of conservation concern such as Atlantic salmon *Salmo salar* (Moore *et al.* 2014; Rougemont & Bernatchez 2018).

Future marine conservation planning is set to be a complex process. Marine life faces universal threats across global oceans while species and populations face their own challenges. Therefore, conservation stands to benefit from combining assessment methods and applying multidisciplinary management approaches tailored to the multiple threat levels experienced. As part of my PhD thesis, I have taken a hierarchal approach to conservation by exploring threats at the global, species and population level. I have undertaken two different methods of applying population genetic methodology to marine sciences and their conservation applications. A synthesis methodology has enabled me to take advantage of the large amount of publicly archived open data, a relatively recent practice that offers compelling research opportunities (Leigh *et al.* 2021; Reichman *et al.* 2011), to uncover global genetic diversity patterns, their drivers and address global conservation questions. I interpret the effects of distinct environmental and human-caused drivers of global genetic diversity to assist efforts to define current biodiversity patterns, status and conservation priorities for marine fish populations, knowledge critical to forecasting future trends and plans for sustainability. In the second approach, I focus on and generate my own data for a single species, in this case anadromous sea lamprey *Petromyzon marinus*, to address species and population level conservation questions. An ancient jawless fish widely distributed on both sides of the North Atlantic Ocean (Potter *et al.*

2015), sea lamprey has potential for wide dispersal during its parasitic feeding phase (Mateus *et al.* 2021) and lacks natal homing behavior typical for anadromous fishes (Bergstedt & Seelye 1995; Waldman *et al.* 2008). These characteristics make it an interesting focus species for population genetic studies. Lack of homing behavior and wide dispersal will minimize noise and enable the detection of adaptation under high gene flow. The wide range of environments sea lamprey inhabits across its distribution range will enable detection of adaptation to local environmental conditions. Increasingly available next generation sequencing methods will be invaluable to uncovering evolutionary patterns and processes underlying anadromous sea lamprey's oceanographically diverse habitat range and life history, while contributing knowledge on evolution of North Atlantic diadromous fishes as a whole. Through this project, I account for the importance of both a single species and global conservation perspective, by simultaneously studying global marine genetic diversity and its drivers while focusing on producing specialized single species knowledge on anadromous sea lamprey. My research seeks to extend understanding of global marine biodiversity status and trends as related to global fish species, sea lamprey population genetic structure, demography and evolutionary history across its distribution range with a conservation and management focus. I interpret results from a conservation genetics point of view, aiming to support efforts for sustainability and inclusion of genetic metrics in marine conservation.

References

Andrews, K.R., Good, J.M., Miller, M.R., Luikart, G. & Hohenlohe, P.A. (2016). Harnessing the power of RADseq for ecological and evolutionary genomics. Nat. Rev. Genet., 17, 81–92.

Barceló, A., Sandoval-Castillo, J., Brauer, C.J., Bilgmann, K., Parra, G.J., Beheregaray, L.B., et al. (2022). Seascape genomics of common dolphins (*Delphinus delphis*) reveals adaptive diversity linked to regional and local oceanography. BMC Ecol. Evol., 22, 88.

Barnosky, A.D., Matzke, N., Tomiya, S., Wogan, G.O.U., Swartz, B., Quental, T.B., et al. (2011). Has the Earth's sixth mass extinction already arrived? Nature, 471, 51–57.

Barrett, R. & Schluter, D. (2008). Adaptation from standing genetic variation. Trends Ecol. Evol., 23, 38–44.

Bergstedt, R.A. & Seelye, J.G. (1995). Evidence for lack of homing by sea lampreys. Trans. Am. Fish. Soc., 124, 235–239.

Bourret, V., Kent, M.P., Primmer, C.R., Vasemägi, A., Karlsson, S., Hindar, K., et al. (2013). SNP-array reveals genome-wide patterns of geographical and potential adaptive divergence across the natural range of Atlantic salmon (*Salmo salar*). Mol. Ecol., 22, 532–551.

Buckley, S.J., Brauer, C.J., Unmack, P.J., Hammer, M.P. & Beheregaray, L.B. (2022). Variation in intraspecific demography drives localised concordance but species-wide discordance in response to past climatic change. BMC Ecol. Evol., 22, 35.

Burt, A.J., Vogt-Vincent, N., Johnson, H., Sendell-Price, A., Kelly, S., Clegg, S.M., et al. (2024). Integration of population genetics with oceanographic models reveals strong connectivity among coral reefs across Seychelles. Sci. Rep., 14, 4936.

Canales-Aguirre, C.B., Larson, W.A., McKinney, G.J., Claure, C.E., Rocha, J.D., Ceballos, S.G., et al. (2022). Neutral and adaptive loci reveal fine-scale population structure in *Eleginops maclovinus* from north Patagonia. Ecol. Evol., 12, e9343.

Candy, J.R., Campbell, N.R., Grinnell, M.H., Beacham, T.D., Larson, W.A. & Narum, S.R. (2015). Population differentiation determined from putative neutral and divergent adaptive genetic markers in Eulachon (*Thaleichthys pacificus*, Osmeridae), an anadromous Pacific smelt. Mol. Ecol. Resour., 15, 1421–1434.

Clemens, B.J., Wyss, L., McCoun, R., Courter, I., Schwabe, L., Peery, C., et al. (2017). Temporal genetic population structure and interannual variation in migration behavior of Pacific Lamprey *Entosphenus tridentatus*. Hydrobiologia, 794, 223–240.

Costello, C., Cao, L., Gelcich, S., Cisneros-Mata, M.Á., Free, C.M., Froehlich, H.E., et al. (2020). The future of food from the sea. Nature, 588, 95–100.

Dahle, G., Johansen, T., Westgaard, J.-I., Aglen, A. & Glover, K.A. (2018). Genetic management of mixed-stock fisheries "real-time": The case of the largest remaining cod fishery operating in the Atlantic in 2007–2017. Fish. Res., 205, 77–85.

De Lafontaine, G., Napier, J.D., Petit, R.J. & Hu, F.S. (2018). Invoking adaptation to decipher the genetic legacy of past climate change. Ecology, 99, 1530–1546.

Ellegren, H. (2014). Genome sequencing and population genomics in non-model organisms. Trends Ecol. Evol., 29, 51–63.

FAO (2022). Towards blue transformation. The state of world fisheries and aquaculture. FAO, Rome.

Gracia, A., Rangel-Buitrago, N., Oakley, J.A. & Williams, A.T. (2018). Use of ecosystems in coastal erosion management. Ocean Coast. Manag., 156, 277–289.

Guannel, G., Arkema, K., Ruggiero, P. & Verutes, G. (2016). The power of three: Coral reefs, seagrasses and mangroves protect coastal regions and increase their resilience. PLOS ONE, 11, e0158094.

Hand, B.K., Flint, C.G., Frissell, C.A., Muhlfeld, C.C., Devlin, S.P., Kennedy, B.P., et al. (2018). A social–ecological perspective for riverscape management in the Columbia River Basin. Front. Ecol. Environ., 16.

Hare, M.P., Nunney, L., Schwartz, M.K., Ruzzante, D.E., Burford, M., Waples, R.S., et al. (2011). Understanding and estimating effective population size for practical application in marine species management: Applying effective population size estimates to marine species management. Conserv. Biol., 25, 438–449.

Hauser, L. & Carvalho, G.R. (2008). Paradigm shifts in marine fisheries genetics: Ugly hypotheses slain by beautiful facts. Fish Fish., 9, 333–362.

Heinze, C., Meyer, S., Goris, N., Anderson, L., Steinfeldt, R., Chang, N., et al. (2015). The ocean carbon sink – impacts, vulnerabilities and challenges. Earth Syst. Dyn., 6, 327–358.

Hemmer-Hansen, J., Hüssy, K., Baktoft, H., Huwer, B., Bekkevold, D., Haslob, H., et al. (2019). Genetic analyses reveal complex dynamics within a marine fish management area. Evol. Appl., 12, 830–844.

Hess, J.E., Campbell, N.R., Close, D.A., Docker, M.F. & Narum, S.R. (2013). Population genomics of Pacific lamprey: Adaptive variation in a highly dispersive species. Mol. Ecol., 22, 2898–2916.

Hess, J.E., Delomas, T.A., Jackson, A.D., Kosinski, M.J., Moser, M.L., Porter, L.L., et al. (2022). Pacific lamprey translocations to the Snake River boost abundance of all life stages. Trans. Am. Fish. Soc., 151, 263–296.

Hess, J.E., Smith, J.J., Timoshevskaya, N., Baker, C., Caudill, C.C., Graves, D., et al. (2020). Genomic islands of divergence infer a phenotypic landscape in Pacific lamprey. Mol. Ecol., 29, 3841–3856.

Hilborn, R., Quinn, T.P., Schindler, D.E. & Rogers, D.E. (2003). Biocomplexity and fisheries sustainability. Proc. Natl. Acad. Sci., 100, 6564–6568.

Hoban, S., Archer, F.I., Bertola, L.D., Bragg, J.G., Breed, M.F., Bruford, M.W., et al. (2022). Global genetic diversity status and trends: towards a suite of Essential Biodiversity Variables (EBVs) for genetic composition. Biol. Rev., 97, 1511–1538.

Hoban, S., Bruford, M., D' Urban Jackson, J., Lopes-Fernandes, M., Heuertz, M., Hohenlohe, P.A., et al. (2020). Genetic diversity targets and indicators in the CBD post-2020 Global Biodiversity Framework must be improved. Biol. Conserv., 248, 108654.

Hughes, A.C. & Grumbine, R.E. (2023). The Kunming-Montreal Global Biodiversity Framework: What it does and does not do, and how to improve it. Front. Environ. Sci., 11, 1281536.

IUCN (2012). IUCN Red List categories and criteria, version 3.1, second edition. 2nd edn. Switzerland and Cambridge, UK: IUCN.

Jouffray, J.-B., Blasiak, R., Norström, A.V., Österblom, H. & Nyström, M. (2020). The blue acceleration: The trajectory of human expansion into the ocean. One Earth, 2, 43–54.

Karachaliou, E., Schmidt, C., De Greef, E., Docker, M.F. & Garroway, C.J. (2025). Urbanization is associated with reduced genetic diversity in marine fish populations. Mol. Ecol., e17711.

Layton, K.K.S., Snelgrove, P.V.R., Dempson, J.B., Kess, T., Lehnert, S.J., Bentzen, P., et al. (2021). Genomic evidence of past and future climate-linked loss in a migratory Arctic fish. Nat. Clim. Change, 11, 158–165.

Lehnert, S.J., Kess, T., Bentzen, P., Kent, M.P., Lien, S., Gilbey, J., et al. (2019). Genomic signatures and correlates of widespread population declines in salmon. Nat. Commun., 10, 2996.

Leigh, D., Vandergast, A., Hunter, M., Crandall, E., Funk, W., Garroway, C., et al. (2023). Don't make genetic data disposable: Best practices for genetic and genomic data archiving.

Leigh, D.M., Hendry, A.P., Vázquez-Domínguez, E. & Friesen, V.L. (2019). Estimated six per cent loss of genetic variation in wild populations since the industrial revolution. Evol. Appl., 12, 1505–1512.

Leigh, D.M., Van Rees, C.B., Millette, K.L., Breed, M.F., Schmidt, C., Bertola, L.D., et al. (2021). Opportunities and challenges of macrogenetic studies. Nat. Rev. Genet., 22, 791–807.

Limborg, M.T., Helyar, S.J., De Bruyn, M., Taylor, M.I., Nielsen, E.E., Ogden, R., et al. (2012). Environmental selection on transcriptome-derived SNPs in a high gene flow marine fish, the Atlantic herring (*Clupea harengus*). Mol. Ecol., 21, 3686–3703.

Longo, G.C., Lam, L., Basnett, B., Samhouri, J., Hamilton, S., Andrews, K., et al. (2020). Strong population differentiation in lingcod (*Ophiodon elongatus*) is driven by a small portion of the genome. Evol. Appl., 13, 2536–2554.

Maeda, Y., Yoshino, T., Matsunaga, T., Matsumoto, M. & Tanaka, T. (2018). Marine microalgae for production of biofuels and chemicals. Curr. Opin. Biotechnol., 50, 111–120.

Mateus, C.S., Docker, M.F., Evanno, G., Hess, J.E., Hume, J.B., Oliveira, I.C., et al. (2021). Population structure in anadromous lampreys: Patterns and processes. J. Gt. Lakes Res., 47, S38–S58.

Milano, I., Babbucci, M., Cariani, A., Atanassova, M., Bekkevold, D., Carvalho, G.R., et al. (2014). Outlier SNP markers reveal fine-scale genetic structuring across European hake populations (*Merluccius merluccius*). Mol. Ecol., 23, 118–135.

Moore, J., Bourret, V., Dionne, M., Bradbury, I., O'Reilly, P., Kent, M., et al. (2014). Conservation genomics of anadromous Atlantic salmon across its North American range: Outlier loci identify the same patterns of population structure as neutral loci. Mol. Ecol., 23, 5680–5697.

Muller-Karger, F.E., Miloslavich, P., Bax, N.J., Simmons, S., Costello, M.J., Sousa Pinto, I., et al. (2018). Advancing marine biological observations and data requirements of the complementary Essential Ocean Variables (EOVs) and Essential Biodiversity Variables (EBVs) frameworks. Front. Mar. Sci., 5, 211.

Neely, K.L., Lewis, C.L., O'Neil, K., Woodley, C.M., Moore, J., Ransom, Z., et al. (2021). Saving the last unicorns: The genetic rescue of Florida's pillar corals. Front. Mar. Sci., 8, 657429.

Palsboll, P., Berube, M. & Allendorf, F. (2007). Identification of management units using population genetic data. Trends Ecol. Evol., 22, 11–16.

Pinsky, M.L. & Palumbi, S.R. (2014). Meta-analysis reveals lower genetic diversity in overfished populations. Mol. Ecol., 23, 29–39.

Potter, I.C., Gill, H.S., Renaud, C.B. & Haoucher, D. (2015). The taxonomy, phylogeny, and distribution of lampreys. In: Lampreys: Biology, conservation and control (ed. Docker, M.F.). Springer Netherlands, Dordrecht, pp. 35–73.

Reichman, O.J., Jones, M.B. & Schildhauer, M.P. (2011). Challenges and opportunities of open data in ecology. Science, 331, 703–705.

Reiss, H., Hoarau, G., Dickey-Collas, M. & Wolff, W.J. (2009). Genetic population structure of marine fish: Mismatch between biological and fisheries management units. Fish Fish., 10, 361–395.

Rougemont, Q. & Bernatchez, L. (2018). The demographic history of Atlantic salmon (*Salmo salar*) across its distribution range reconstructed from approximate Bayesian computations: Atlantic salmon history and linked selection. Evolution, 72, 1261–1277.

Sala, E. & Knowlton, N. (2006). Global marine biodiversity trends. Annu. Rev. Environ. Resour., 31, 93–122.

Schmidt, C., Hoban, S., Hunter, M., Paz-Vinas, I. & Garroway, C.J. (2023). Genetic diversity and IUCN Red List status. Conserv. Biol., 37, e14064.

Spies, I. & Punt, A.E. (2015). The utility of genetics in marine fisheries management: A simulation study based on Pacific cod off Alaska. Can. J. Fish. Aquat. Sci., 72, 1415–1432.

Tennakoon, P., Chandika, P., Yi, M. & Jung, W.-K. (2023). Marine-derived biopolymers as potential bioplastics, an eco-friendly alternative. iScience, 26, 106404.

Van Wyngaarden, M., Snelgrove, P.V.R., DiBacco, C., Hamilton, L.C., Rodríguez-Ezpeleta, N., Zhan, L., et al. (2018). Oceanographic variation influences spatial genomic structure in the sea scallop, *Placopecten magellanicus*. Ecol. Evol., 8, 2824–2841.

Waldman, J., Grunwald, C. & Wirgin, I. (2008). Sea lamprey *Petromyzon marinus*: An exception to the rule of homing in anadromous fishes. Biol. Lett., 4, 659–662.

Xuereb, A., D'Aloia, C.C., Daigle, R.M., Andrello, M., Dalongeville, A., Manel, S., et al. (2019). Marine conservation and marine protected areas. In: Population genomics: Marine organisms, population genomics (eds. Oleksiak, M.F. & Rajora, O.P.). Springer International Publishing, Cham, pp. 423–446.

CHAPTER 2: HUMAN-CAUSED AND ENVIRONMETAL DRIVERS OF GLOBAL GENETIC DIVERSITY IN MARINE FISH SPECIES

Karachaliou E. [1], Schmidt C. [2], de Greef E. [1], Müller C. [1], Docker M.F. [1], Garroway C.J. [1]

[1] Department of Biological Sciences, University of Manitoba, 50 Sifton Rd, Winnipeg, MB, R3T 2N2, Canada

[2] Department of Biology, Dalhousie University; 1355 Oxford St, Halifax, NS, B3H 3Z1, Canada

Results on human population density, cumulative human impacts and fishing pressure published as a standalone study (https://onlinelibrary.wiley.com/doi/full/10.1111/mec.17711).

Abstract

As humans drive biodiversity declines across global oceans, maintaining genetic diversity is essential as it underlies adaptive potential and overall population persistence. A first step to conserving genetic diversity in marine fishes is understanding global spatial patterns, the oceanographic processes that shape genetic diversity and relationship with species richness. I also examine if human activity is causing declines in genetic diversity. I programmatically searched open data repositories collecting, georeferencing and re-purposing raw neutral microsatellite data for marine fishes – 75,496 individuals, 1,143 populations and 73 species. I then tested for relationships between genetic diversity and primary productivity, distance to market, nearby land area, species richness, human population density, fishing effort and cumulative human impacts on genetic diversity. I uncover global spatial patterns in marine fishes' genetic diversity, oceanographic drivers, human caused declines, and the relationship between species richness and genetic diversity. I then discuss conservation implications.

Introduction

Human behavior is contributing to biodiversity declines by extracting resources, polluting, and altering habitats across global oceans (Sala & Knowlton 2006). The detrimental effects of overexploitation and subsequent societal implications have been extensively reviewed in scientific literature, along with proposed solutions. (Berkes *et al.* 2006; Duarte *et al.* 2020; Worm *et al.* 2006, 2009). Genetic diversity facilitates population persistence as it underlies adaptive potential and ecosystem stability and resilience (Raffard *et al.* 2019) in the face of climate change (Barrett & Schluter 2008). Existing genetic variation has been known to facilitate evolutionary responses to the effects of climate change in the marine environment, for example, resilience to higher temperatures found in corals (Bay & Palumbi 2014) and ocean acidification in sea urchin (Pespeni *et al.* 2013). Global efforts for genetic diversity conservation are underway, with one of the Kunming-Montreal Biodiversity Framework long-term goals for 2050 being to maintain genetic diversity within populations across species and safeguard adaptive potential. If we are to realize these goals, it is important to identify to what extent centuries of human activity have affected genetic diversity across global oceans and uncover underlying spatial patterns and associations between local environment conditions and genetic diversity. Connecting underlying environmental conditions to evolutionary processes and accounting for the effect of human disturbance could help predict evolutionary responses to future changing conditions, assess vulnerability and support a science-based approach to safeguarding biodiversity (e.g., Hoban *et al.* 2022; Muller-Karger *et al.* 2018).

Existing studies have shown varying genetic diversity trends across different markers, taxa and environments. No spatial patterns of neutral nuclear genetic diversity have been found for marine fishes, but genetic diversity is associated with primary productivity (Clark & Pinsky

2024). Evidence of human impacts at the genetic level has already been detected. A meta-analysis by Pinsky & Palumbi (2014) found that genetic diversity is lower in marine fishes subject to overharvesting. Meanwhile on land, continued human encroachment on wild habitat has come under scrutiny. Urbanization has been found to negatively affect mammal genetic diversity in North America (Schmidt *et al.* 2020), but no overarching relationships were found for either amphibians (Schmidt & Garroway 2021b) or birds (Schmidt *et al.* 2020). Even problematic human ideologies have been found to leave a lasting effect on wildlife. Government-imposed residential racial segregation policies in U.S. cities have led to genetic diversity declines and fragmentation in urban wildlife (Schmidt & Garroway 2022). Understanding environmental processes underlying genetic diversity patterns and their spatial distribution will help inform conservation planning and predict how environmental changes linked to climate change and future population declines in wild species may shape global marine genetic diversity.

To date, literature on genetic diversity patterns and drivers has mainly focused on mitochondrial DNA markers. Global mitochondrial genetic diversity is connected to species richness patterns (Manel *et al.* 2020), primary productivity (Clark & Pinsky 2024) and positively related to sea surface temperature (Clark & Pinsky 2024; Manel *et al.* 2020) for marine fish species. Regions of high diversity are found in the West Pacific, North Indian Ocean and the Caribbean, while regions with low diversity are found in the North-Eastern and West Atlantic and the South Atlantic oceans (Manel *et al.* 2020; Clark & Pinsky 2024). At the local level, mitochondrial genetic diversity has been found to differ across South African bioregions for marine fish and invertebrate species (Dalongeville *et al.* 2022). Although informative for addressing some species-specific conservation questions, mitochondrial DNA markers are not advised for conservation applications in a global context as they do not capture ecology,

demography and genome-wide diversity (Schmidt & Garroway 2021a). Therefore, my study aims to uncover patterns and drivers of genetic diversity in global marine fish species based on microsatellite markers, a good indicator of genome-wide diversity (Mittell *et al.* 2015).

Species richness, the best-studied biodiversity pattern, displays a latitudinal gradient in coastal fish species, with diversity highest in the tropical Indo-Pacific and small bimodal peaks north and south of the equator. Low to mid-latitude regions of high diversity were observed in pelagic fish species, and slightly higher levels were observed in the southern hemisphere (Worm & Tittensor 2018b). Species richness was also positively related to sea surface temperature across all species, the effects of habitat area in coastal species, and habitat complexity in pelagic species (Worm & Tittensor 2018a). A connection between nuclear genetic diversity and species richness has yet to be made for marine fishes. Uncovering their relationship will be an asset to conservation planning as different approaches will be needed if these two metrics do not follow similar patterns. In the marine realm, a positive association between marine habitat-forming species richness and nuclear genetic diversity has been found within ecoregions (Figuerola-Ferrando *et al.* 2023). Meanwhile on land, nuclear genetic diversity and species richness show a negative relationship and can be explained by resource availability and ecological opportunity in North American mammals (Schmidt *et al.* 2022a). In North American amphibians, species richness and genetic diversity are also negatively correlated but high variation in life history makes finding environmental correlates of genetic diversity and species richness difficult (Schmidt *et al.* 2022b).

In this study, I synthesize publicly available nuclear microsatellite data sets archived with individual studies to map genetic diversity in marine fish populations and explore natural and human-caused correlates of genetic diversity. Microsatellites are well correlated with genome-

wide diversity (Mittell *et al.* 2015) and are the most abundant and best spatially sampled marker type in data archives. Markers with genome-wide coverage such as SNPs are not as readily available in data archives, therefore microsatellites were preferred for the needs of this study. Isolating environmental and human-caused drivers will help define the patterns and processes underlying marine fish genetic diversity and inform conservation efforts. I predict that human activity will negatively affect genetic diversity and connectivity. To explore this, I use distance to market (Yeager *et al.* 2017), fishing pressure (Kroodsma *et al.* 2018), human population density (CIESIN 2018b) and cumulative human impacts (Halpern *et al.* 2015) as metrics for potential human-caused declines in genetic diversity. I also predict that genetic diversity and connectivity will be correlated with variations in local oceanographic conditions. Here I test for associations with primary productivity (Yeager *et al.* 2017), dissolved oxygen (Assis *et al.* 2018; Tyberghein *et al.* 2012), pH (Assis *et al.* 2018; Tyberghein *et al.* 2012), salinity (Assis *et al.* 2018; Tyberghein *et al.* 2012), sea surface temperature (Assis *et al.* 2018; Tyberghein *et al.* 2012) and land area (CIESIN 2018a). I also seek to understand the connection between biodiversity at the species level, species richness, and intraspecific genetic diversity. Results are interpreted in regard to both current and future conditions. I consider how the potential effects of human activity may translate in the future under amplified human impacts, climate change and a growing human population. I also consider implications of potential variation in genetic diversity found across environmental gradients and climate change-derived environmental alterations such as rising temperatures, ocean acidification, ocean deoxygenation and increased instances of eutrophication.

Methods

Data compilation

My data collection and analyses follow Schmidt & Garroway (2021b) and Schmidt *et al.* (2020). To compile data sets, I programmatically searched online publicly accessible data repositories with the R DataONE interface (Jones *et al.* 2016) using the keywords "microsat," "microsatellite," and species name (e.g., "*Petromyzon marinus*"). A species list (8506 species) was downloaded from the IUCN Red List database (www.iucnredlist.org) using class ("Actinopterygii", "Chondrichthyes", "Petromyzontida") and marine regions of interest as search terms. I included diadromous and brackish species inhabiting transitional zones. After assessing suitability, I was left with 84 usable search results (Supplemental Information 1.1). Datasets were considered usable when they consisted of presumed neutral microsatellites (e.g., excluding known adaptive markers and SNPs) generated in a way that would not affect genetic diversity (e.g., excluding individuals reared in captivity and simulated data). If datasets lacked corresponding coordinates but included a sample map, coordinates were produced by georeferencing in QGIS version 3.16 (QGIS Development Team 2021). As the scope of this study does not include historical data, a cut-off point for usable samples was set to the year 1990. Populations known to have been subject to restocking practices and populations post-natural disaster were included in the study, as they are considered to reflect the present impacted state of these populations. After applying these dataset retention criteria, I had 108 usable microsatellite datasets (Supplemental Information 1.2).

These raw microsatellite datasets were used to calculate gene diversity, allelic richness, population-specific F_{ST} and effective population size for each population in R version 4.1.2 (R Core Team 2021). Gene diversity (Nei 1973), a measure of the probability that two alleles randomly selected from the population are the same, is largely unaffected by sample size

(Charlesworth & Charlesworth 2010) and was estimated using the "Hs" function in the adegenet package in R (Jombart 2008; Jombart & Ahmed 2011). Allelic richness, an allele count, is sensitive to how many individuals are sampled (Leberg 2002) and was calculated using the "allelic.richness" function in the hierfstat package (Goudet & Jombart 2015) for populations with minimum number of alleles = 10. Allelic richness is affected by effective population size and bottlenecks and therefore best used for understanding contemporary evolutionary processes (Luikart *et al.* 1998) and admixture (Chakraborty *et al.* 1988). Gene diversity is less sensitive to these influences compared to allelic richness (Nei *et al.* 1975) and is best used to understand past processes. Population-specific F_{ST} (Weir & Goudet 2017), a measure of between-population genetic differentiation relative to an assumed common ancestor, was calculated using the "betas" function from hierfstat, relying on a minimum of two populations per dataset. Effective population size, or the rate at which a population is susceptible to genetic drift, was calculated based on a linkage disequilibrium method using NeEstimator (Do *et al.* 2014). Effective population size could not be calculated at all sites due to signals of genetic drift being lost to sampling error. I log-transformed effective population size for use in further analysis and filtered out populations with fewer than five individuals.

Next, I considered how human actions could affect genetic diversity for marine fish species. Initially, I considered a population's distance to market, interpreted as its likelihood to be a target of the seafood industry, using straight-line distance to provincial capital as a metric of market access from Yeager *et al.* (2017). Human population density can be considered an indicator of urbanization and increased human activity-derived pressures near large human population centers. For this metric, I downloaded estimates of human population density (number of persons per square kilometer) in accordance with information from national censuses

and population registers for the year 2010 in a 30 arc-second resolution as accessed from CIESIN (2018b). To account for the sum of human impacts on the world's oceans, I used the metric cumulative human impacts for the year 2013 from Halpern *et al.* (2015); cumulative human impacts is an indicator of the effects of climate change, fishing, and ocean- and land-based stressors on marine ecosystems. To further isolate the influence of fishing pressure, I used hours of fishing per km^2 (Kroodsma *et al.* 2018) as a metric.

To uncover the environmental drivers of genetic diversity I used metrics of mean dissolved molecular oxygen ($mmol/m^{-3}$), mean surface pH, mean surface salinity (PSS) and mean sea surface temperature (ºC) downloaded from Bio-ORACLE (Assis *et al.* 2018; Tyberghein *et al.* 2012). I also used net primary productivity as extracted from 8-day composite layers from 2003-2013 produced by NOAA CoastWatch (Yeager *et al.* 2017). To account for the influence of nutrient and energy runoff from terrestrial to associated marine ecosystems, I downloaded estimates of land area (square kilometers per pixel) in a 30 arc-second resolution as accessed from CIESIN (2018a). I obtained marine fish species richness estimates from Jenkins & Van Houtan (2016). To account for effects across local to wider-scale areas, buffers were applied when extracting information from these datasets at a 25, 50, 100 and 200 km radius. Datasets were re-projected to the Lambert Cylindrical Equal Area coordinate system to assure they are compatible and comparable.

Estimating spatial patterns of genetic diversity

To visualize patterns of spatial variation in genetic diversity, I created maps of gene diversity, allelic richness, population-specific F_{ST} and Ne, following methodology described in Schmidt *et al.* (2022a), in R version 3.6.3 (R Core Team 2020). I used distance-based Moran's eigenvector maps (MEMs) (Dray *et al.* 2006), orthogonal spatially explicit eigenvectors that

account for spatial autocorrelation across all scales with eigenvalues proportional to Moran's I. MEMs were calculated in the R package adespatial (Dray *et al.* 2017). I used the R package marmap (Pante & Simon-Bouhet 2013) to create a distance matrix of spatial connections suitable for a global marine dataset. Forward selection (Blanchet *et al.* 2008) was used to select MEMs representing significant patterns in gene diversity, allelic richness, population-specific F_{ST} and effective population size. MEMs representing broad spatial patterns in my data were selected based on a threshold of Moran's $I > 0.25$. I then created maps of gene diversity, allelic richness and population-specific F_{ST} using predictions from MEM models associated with the genetic metric of interest. In the case of effective population size, where no significant MEMs could be found, I used raw values for visualization.

Statistical analyses

I performed the following analyses in R version 3.6.3 (R Core Team 2020). I tested for the effect of human behavior and surrounding environmental parameters on genetic diversity using Bayesian generalized linear mixed models with the brms package (Bürkner 2017). Predictor and response variables were scaled and centered before running models. Species were treated as a random effect and between-species variation in the means of each genetic metric were controlled for using random intercepts. Differences in relationships with predictors were controlled for with slopes permitted to vary with species. In summary, I ran models for each candidate measure of genetic diversity as explained by each candidate human-caused/environmental driver including species as a random effect. All models in this study were run with four chains and a minimum of 2,000 iterations.

In total, I ran 176 models: 4 genetic diversity measures x 11 candidate environmental or human-caused drivers x 4 radii. Models were run both with and without priors, and I used normal

priors with a mean 0 and a standard deviation 1 [Normal (0,1)] from the regression parameter representing the corresponding population-level effect. Priors were conservative and accounted for multiple testing issues. I tested for spatial autocorrelation in the model residuals using Moran's I test in the package spdep (Bivand & Wong 2018). No spatial autocorrelation was found, so I proceeded with my models without spatial corrections.

I estimated the probability of an effect being in the direction diagnosed by the 95% credible intervals by calculating probability of direction of effect measured with the function "p_direction" in bayestestR (Makowski *et al.* 2019). This summarizes the proportion of values in the posterior distribution that have the same sign as the median value. Models with >89% (McElreath 2018) probability of direction were considered to be of significant effect. I used the brms package to perform posterior predictive checks using the function pp_check. Further model diagnostics to assess model performance, efficient approximate leave-one-out cross-validation (LOO) and widely applicable information criterion (WAIC) were estimated using loo() and waic() respectively from the brms package. Probability of direction, LOO and WAIC were used to select the most significant spatial scale per model. Conditional and marginal R^2 values were estimated using r2_bayes() in the package performance (Lüdecke *et al.* 2021). Conditional R^2 is the variance explained by both fixed and random effects while marginal R^2 is the variance explained only by the fixed effects. Bayesian R-squared values were estimated using the function bayes_R2() in brms. Bayesian R^2 is the variance explained by both fixed and random effects divided by the total variance in the data. I tested for correlation between candidate drivers by computing a correlation matrix with the package ggcorrplot (Kassambara 2023).

Results

My initial global ocean dataset consists of 75,496 individuals, 1143 populations and 73 fish species (Figure 2.1, Supplemental Information 1.2, 1.3). Based on data availability, I successfully calculated gene diversity and allelic richness for 1143 populations. Per population F_{ST} was calculated for 1131 populations, as datasets with less than two populations were excluded. Effective population size was calculated for 682 populations. The signal of genetic drift can be lost due to sampling error and Ne estimated as infinity. I excluded these estimates from further analysis. Maps of raw genetic data are available in Figure 2.5, Supplemental Information 1.4-1.6. Populations with fewer than five individuals were removed, leaving 1085 populations, 75,361 individuals and 73 species for subsequent analysis. The dataset I compiled, aside from the purposes of this thesis and planned future publications, has been used as part of an additional two studies I am co-author on (Paz-Vinas *et al.* 2023, Schmidt *et al.* 2024).

Global spatial patterns

Spatial variation followed the same pattern for gene diversity and allelic richness (Figures 2.2 & 2.3). High levels of gene diversity and allelic richness were found in the North Atlantic, Baltic, Mediterranean Sea, Gulf of Mexico and Caribbean regions. Intermediate values were found in the South Atlantic, Indian Ocean, West and Central South Pacific Oceans and the Red Sea, with a decline along a west to east gradient. Low levels of gene diversity and allelic richness were found in the East and Central North Pacific Ocean. F_{ST} followed the opposite pattern (Figure 2.4), low levels found in the North Atlantic, Baltic, Mediterranean Sea, Gulf of Mexico and Caribbean regions. Intermediate values were found in the South Atlantic, Indian Ocean, West and Central South Pacific Oceans and the Red Sea, increasing along a west to east gradient. High levels of F_{ST} were found in the East and Central North Pacific Ocean. No patterns of spatial variation were found for Ne, allowing it to be visualized by plotting the raw data (Figure 2.5).

Environmental and human-caused associations

Model effects were similar whether priors were used or not, and I interpret results based on models with priors. The most significant model across all spatial scales examined was then selected for each candidate driver of genetic diversity based on probability of direction, LOO and WAIC (Table 2.1 & Figure 2.6). Diagnostics for models across all buffers, with and without priors can be found in Supplemental Information 1.7, 1.8 respectively. Plots for models across all buffers, with and without priors can be found in Supplemental Information 1.9-1.40. When testing for correlation between candidate drivers of genetic diversity, temperature was found to be positively correlated with species richness and negatively correlated with dissolved oxygen.

I found human population density to be associated with lower gene diversity, allelic richness and effective population size. I found no relationship between human population density and F_{ST}. Cumulative human impacts are associated with lower allelic richness. I found no relationship between cumulative human impacts and gene diversity, F_{ST} or effective population size. I found fishing pressure to be negatively correlated with gene diversity. I found no relationship between fishing pressure and allelic richness, F_{ST} or effective population size. I have published my results on human population density, cumulative human impacts and fishing pressure as a standalone study (Karachaliou et al. 2025). Distance to market is associated with lower gene diversity and higher F_{ST}. I found no relationship between distance to market and allelic richness, or effective population size.

Primary productivity is associated with higher gene diversity and allelic richness, and lower F_{ST}. I found no relationship between primary productivity and effective population size. Land area is associated with higher gene diversity and allelic richness and lower F_{ST}. I found no relationship between land area and effective population size. I found sea surface temperature to be associated with higher allelic richness and lower F_{ST}. I found no relationship between

temperature and gene diversity or effective population size. I found dissolved oxygen to be associated with lower allelic richness and higher F_{ST}. I found no relationship between oxygen and gene diversity or effective population size. Salinity and pH were not associated with any genetic diversity measures. I found species richness to be positively correlated with within species allelic richness and negatively correlated with F_{ST} and effective population size. I found no relationship between species richness and gene diversity.

Discussion

Nuclear genetic diversity patterns differ from mitochondrial genetic diversity and species richness patterns highlighting the differences between these concepts of biodiversity in marine fishes. Genetic diversity and differentiation display opposing spatial patterns; higher between population gene flow leads to higher genetic diversity as it likely facilitates higher rates of evolution.

Global spatial patterns

My study found that spatial patterns of global nuclear genetic diversity differed from those detected for biodiversity at the pelagic/coastal species level in Worm & Tittensor (2018b) and mtDNA diversity (Manel *et al.* 2020). A notable exception is the Caribbean, a secondary location for high coastal fish species richness, where estimates for both nuclear and mitochondrial genetic diversity are high (Manel *et al.* 2020; Worm & Tittensor 2018b). Another exception is secondarily high pelagic fish species richness in the Atlantic thought to be associated with the Gulf Stream and other frontal systems (Worm & Tittensor 2018b). Despite these two exceptions at the local scale, my results highlight the differences between nuclear

genetic diversity, species richness and mitochondrial genetic diversity as metrics of global biodiversity.

Higher allelic richness and gene diversity in localities with low F_{ST}, and low allelic richness and gene diversity in localities with high F_{ST}, suggest that higher connectivity and thus similarity between populations is connected to higher genetic diversity. Opposing patterns between global genetic diversity and genetic differentiation mean that high levels of gene flow are correlated with high diversity in marine fish species. Population connectivity and gene flow could be driving faster evolution as more genetic information is available to evolve from, leading to higher genetic diversity. It is expected that higher connectivity between populations leads to higher genetic diversity (reviewed in Kahilainen *et al.* 2014) and high rates of gene flow are associated with higher levels of genetic diversity at the species level (e.g., Consuegra *et al.* 2005). Low gene flow due to habitat fragmentation has been connected to lower genetic diversity in human-impacted environments (Hoffmann *et al.* 2015; Johnson & Munshi-South 2017), but this principle could also apply to natural environments fragmented by local oceanographic conditions and natural barriers.

The geographic gradient in genetic diversity observed, seemingly following a west to east path along global coastlines, could imply regional causes. Between-basin differences in biophysical conditions, such as oceanographic parameters and human intervention, could be driving differences in genetic diversity locally. These between-ocean differences in genetic diversity suggest that ecosystems can be functional at different genetic diversity levels. Further study of ecosystems and regions successfully functioning at lower genetic diversity levels could assist managing genetic diversity declines in areas with historically higher genetic diversity. Detecting and isolating human influence could inform decisions to manage the effects of human

activity. For example, by finding a connection between urbanization and lower genetic diversity at the global scale my results encourage sustainable urban planning and resource use to help preserve coastal biodiversity. Similar studies could be undertaken locally. Understanding the processes driving genetic diversity at the local level can inform biodiversity conservation efforts locally, and by extension contributing to conserving global biodiversity.

Spatial patterns of genetic diversity could also reflect evolutionary and climatic processes (Bowen *et al.* 2016; Hewitt 2000). Areas of high genetic differentiation detected across the Indo-Pacific and Red Sea align with regional biogeographic barriers (Bowen *et al.* 2016) that combined with large oceanic distances could explain the highly differentiated populations and low levels of genetic diversity detected. Opposing patterns of genetic diversity and differentiation in the north and central Atlantic could result from the isolation of the tropical Atlantic from the Indo-Pacific by historical geological events (reviewed in Bowen *et al.* 2016). Additionally, existing contact between the south Atlantic and Indian Ocean via the complex South African oceanographic system (Bowen *et al.* 2016) could explain the genetic diversity gradient observed, though limited data availability in the region prohibits too much speculation. The lack of data for the Atlantic and Pacific coasts of South America also prevents fully understanding evolutionary processes in the region. Further insight could be gained by examining per ocean patterns, but this is outside of the scope of this study.

Urbanization is connected to lower genetic diversity and higher vulnerability to loss of genetic diversity but not population fragmentation in marine fishes, highlighting humans as a contemporary evolutionary force in wild populations.

Urbanization

Human population density, a measure of urbanization, was associated with lower gene diversity and allelic richness. Fish populations inhabiting environments adjacent to large human population centers are more likely to have low gene diversity and allelic richness. Human population density is also associated with reduced effective population size, indicating that populations associated with large human population centers are more susceptible to genetic drift than those inhabiting less disturbed environments. Genetic drift is thought to be stronger in urbanized terrestrial environments (Johnson & Munshi-South 2017), and my results extend this theory to the marine realm. Therefore, vulnerability to loss of genetic diversity increases in more urbanized marine environments. Marine environments undergo biodiversity changes alongside urbanization (Todd *et al.* 2019), while genetic diversity declines are expected in small populations as a result of habitat fragmentation (Johnson & Munshi-South 2017).

I found no urbanization driven disruption of gene flow. Genetic diversity declines in urbanized marine environments are most likely related to direct disturbances such as population declines, habitat loss and pollution rather than a result of habitat fragmentation. Reviewed in Todd *et al.* (2019), urbanization is a strong selective pressure (Alberti 2015; Donihue & Lambert 2015) leading to population-level adaptation (Alberti *et al.* 2017; Johnson & Munshi- South 2017). Although my results are unable to measure adaptation directly, they show that global fish populations are indeed subject to evolutionary forces caused by continuous human encroachment on marine habitats. The ability to adapt to these pressures will be inhibited by low-standing genetic diversity and increased vulnerability to the effects of genetic drift in highly urbanized environments. We should therefore not expect adaptation to greatly contribute to population persistence in human impacted environments.

A low but detectable negative relationship between genetic diversity and fishing pressure was found, cautioning against unsustainable fishing practices.

Fishing

I found fishing pressure to be negatively correlated with gene diversity, indicating that gene diversity is lower in areas of high fishing activity. Overfishing has been connected to declines in genetic diversity (Pinsky & Palumbi 2014). While effects across global fish species are not as strong, the legacy of human fishing practices is still evident. The joint negative influence of both urbanization, and to a lesser extent fishing, supports the relationship between urbanization and resource extraction (Kirby 2004; Li 2003; reviewed in Todd *et al.* 2019). My results could reflect the high intensity fishing activity taking place coastally, within the reach of human population centers. However, fish populations inhabiting the high seas, defined as marine areas beyond national jurisdiction, were underrepresented in this study so the full extent of human impacts on genetic diversity in the open ocean is yet to be understood. This uncertainty provides further reason to not fish the high seas (Sumaila *et al.* 2015).

The low but detectable negative influence of the sum of human impacts across global oceans on genetic diversity highlights the implications of human encroachment on marine ecosystems.

Human impacts

Cumulative human impacts (Halpern *et al.* 2015), a peer-reviewed index of the sum of human impacts on the world's oceans, are associated with lower allelic richness, implying that the sum of human-caused stressors in the marine environment is associated with lower allelic richness compared to more “pristine” environments. Together with human population density, my results show overall lower genetic diversity in environments associated with higher human

presence. The weaker effect found for the cumulative human impacts metric is likely due to its inclusion of spatiotemporally varying factors such as fishing pressure and sea surface temperature anomalies. These factors vary across space and time compared to the effect of large urban centers that is spatially fixed and persistent across time.

Distance to market is not a good predictor of exploitation in marine fishes at the global level.

Distance to market

Distance to market is negatively correlated with gene diversity; the further a population is from a provincial capital, the lower gene diversity it harbors. Distance to market was also associated with higher F_{ST}; the further a population is from a provincial capital the more genetically divergent it is. I initially considered that distance to market, known to be linked to condition of reef fisheries (Cinner *et al*. 2013), as a predictor of exploitation caused declines in genetic diversity and subsequent fragmentation linked to human proximity. Based on these results, it seems that despite the utility of this metric in more localized studies, for global studies such as this, the intended metric of human activity is not captured. The globalized nature and accessibility of contemporary fish markets and fisheries supply chains could be the reason distance to market is not significant globally. Instead, it seems to be an expression of distance to shore and can be connected to the joint influence of productivity and land area metrics, based on similar results found and discussed below and the fact that productivity and land mass is higher closer to shore. Distance to market cannot be considered a reliable metric for this study and caution is advised when interpreting associated results.

Nutrient rich environments with higher energy availability can likely support larger, more genetically diverse and similar marine fish populations through population persistence and more individuals leading to higher levels of breeding.

Primary productivity

Primary productivity is positively correlated with gene diversity and allelic richness and negatively correlated with F_{ST}. My results show that highly productive environments are connected to increased genetic diversity and similarity. The relationship between genetic diversity and primary productivity is also supported by Clark & Pinsky (2024). Primary productivity can be interpreted as the availability of nutrients or otherwise energy availability. The relationship between genetic diversity and primary productivity could be connected to the energy hypothesis (Currie 1991; Storch *et al*. 2018); regions with higher energy availability may be able to support larger populations with high genetic diversity due to population persistence. Productivity gradients are also known to drive evolution and population fragmentation (Saenz-Agudelo *et al.* 2015), and I consider that this could lead to maintenance of higher genetic similarity within regions of high productivity by limiting gene flow between areas of different productivity levels. Genetic diversity has also been positively connected to population size (McCusker & Bentzen 2010), energy availability potentially being an underlying driver. This could explain increased gene flow and thus similarity within areas of high productivity as more individuals lead to higher rates of breeding.

Land area can be considered a measure of productivity, habitat area and complexity. Marine fish populations with more adjacent land area were found to be more genetically diverse and similar.

Land area

Land area is associated with higher gene diversity and allelic richness and lower F_{ST}. Higher terrestrial nutrient and energy runoff to associated marine ecosystems results in higher genetic diversity and increasingly genetically similar populations. Land area can be considered a productivity metric due to the influence of nutrient runoff and therefore is connected to primary productivity. This is supported by similar relationships found across these two metrics. Coastline length as an expression of available habitat and its complexity is considered a predictor of high biodiversity regions (Worm & Tittensor 2018a) and could also explain the land area metric. Results could also be related to the positive relationship between biodiversity and habitat area (De Souza Júnior *et al.* 2014). High genetic diversity found in proximity to larger land masses is closer to the reach of humans. This is concerning as the continuous expansion of urban settlements could expose more populations to the urbanization-caused genetic diversity declines identified in this study.

Higher temperatures are connected to more genetically diverse and similar marine fish populations. This is likely due to higher temperatures leading to higher metabolic and mutation rates and shorter generation times.

Sea surface temperature

Temperature is connected to ectotherm metabolic processes, species habitat ranges and their biological limits in the marine environment (Deutsch *et al*. 2020, 2024) and is known to underlie evolution and divergence in marine fishes (Gagnaire *et al.* 2012, Riccioni *et al.* 2013, Therkildsen *et al.* 2013). I found temperature to be associated with higher allelic richness. This result supports the expectation of higher genetic diversity under the evolutionary speed

hypotheses, where higher temperatures are thought to drive higher metabolic and mutation rates and shorter generation times (Fine 2015). A similar relationship has also been observed between both mitochondrial genetic diversity and temperature (Manel *et al.* 2020) and biodiversity and temperature (Worm & Tittensor 2018a). I found temperature to be associated with lower F_{ST}. Populations inhabiting environments with higher temperatures are more genetically similar. Based on my results on patterns of global genetic diversity, I expect higher metabolic and mutation rates and shorter generation times under the evolutionary speed hypotheses to drive higher rates of evolution and thus genetic similarity. However, this is contrary to expectations of increased divergence under the evolutionary speed hypotheses.

Lower oxygen environments are associated with more genetically diverse and similar marine fish populations, likely connected to metabolic processes and distribution limits linked to temperature.

Dissolved oxygen

Oxygen is known to influence biodiversity across geological time scales (Sperling *et al.* 2022). Changes in biodiversity are observed across oxygen gradients (Rogers 2000), while driving population structure and divergence (Berg *et al.* 2015; Harniman *et al.* 2013). I found dissolved oxygen to be negatively correlated with allelic richness, areas with lower oxygen levels harboring higher allelic richness. I also found dissolved oxygen to be positively correlated with F_{ST}, meaning that at lower oxygen levels populations are more genetically similar. Oxygen and temperature are intrinsically related in the marine environment, my results reflecting this relationship in connection to genetic diversity. In marine ectotherms, oxygen, alongside temperature, influences metabolic processes and species distribution ranges (Deutsch *et al.* 2020;

2024). Unless drastic measures are taken, rising temperatures and deoxygenation caused by climate change will force species movement and extinction, affecting global biodiversity distribution. The relationship between oxygen, temperature, species richness and projections under global change (Deutsch *et al*. 2024), in connection to what is now known about these parameters' relationship to genetic diversity, urges inclusion of genetic diversity in climate models.

Variation in salinity and pH levels doesn't seem to be relevant to marine fishes' genetic diversity or differentiation at the global level.

Salinity & pH

Salinity was not associated with genetic diversity or differentiation, despite it being a known driver of evolution and differentiation across environmental gradients (Berg *et al.* 2015; Gaggiotti *et al.* 2009; Riccioni *et al.* 2013). A driver of adaptive evolution under simulated future climate change (Pespeni *et al.* 2013), pH was not associated with any genetic diversity measures. There are known cases specific to ocean acidification conditions where adaptation benefits from standing genetic variation (Bitter *et al.* 2019) but is inhibited by stronger effects of short-term genetic drift (Harvey *et al.* 2016). If genetic diversity is to be incorporated in plans to combat ocean acidification, background information on underlying natural processes is needed. If the intention is to maintain high levels of genetic diversity within populations subject to ocean acidification conditions, the lack of a relationship between natural gradients in pH and genetic diversity means that more generalized policies could be applicable at least globally. However, this principle should not be applied locally before studies on finer-scale effects have been done.

Species rich environments lead to more genetically diverse and similar marine fish populations, potentially driven by energy availability, higher metabolic and mutation rates and shorter generation times. Species rich environments seemingly support small populations due to resource competition intensity as evident by low effective population sizes.

Species richness

I found species richness, a metric of species level biodiversity, to be positively correlated with allelic richness within species. Species richness in marine fishes is associated with higher temperature (Worm & Tittensor 2018a), and the positive relationship with allelic richness could be connected to both the evolutionary speed (Fine 2015) and energy (Currie 1991; Storch *et al.* 2018) hypotheses. This positive relationship between genetic diversity and species richness for marine fishes also reflects the results of Manel *et al.* (2020). Species richness was negatively correlated with F_{ST} within species, with higher species richness leading to increasingly genetically similar populations. Species richness is also associated with lower within species effective population size; the more species concentrated in an area, the lower effective population size is and the higher populations' susceptibility to genetic drift. I consider that this negative relationship could be caused by higher resource competition intensity in species rich environments maintaining smaller population sizes with less breeding individuals. Species richness is linked to resource availability and habitat complexity, with lower fitness enabling more species to coexist as resources increase, until a peak is reached, and species richness declines due to competition (De Souza Júnior *et al.* 2014). Results could reflect the genetic implications of loss of fitness due to increased competition for resources leading to smaller populations in species rich environments. Mitochondrial (Manel *et al.* 2020) and nuclear genetic diversity are associated with higher species richness, while all are associated with higher sea

surface temperature (Manel *et al.* 2020; Worm & Tittensor 2018a). My results add nuclear genetic diversity to this relationship and support the interconnectedness of biodiversity and temperature in the marine environment.

I expect both human-caused and environmental parameters to influence future genetic diversity dynamics under climate change.

Implications

Urbanization is associated with lower effective population size, increasing vulnerability to the effect of genetic drift. Evidence that effective population size in marine fishes is negatively affected by humans is evidence against the belief that large, highly connected marine populations would be less sensitive to population declines, and subsequently effects of genetic drift. Subject to genetic drift, genetic variation is eroded reducing fitness and the efficiency of natural selection, and effective population size is a recommended measure of population resilience in management contexts as it takes into account both genetic and life history parameters (reviewed in Hare *et al.* 2011). Observed higher species richness associated with lower Ne reflects natural processes at play, sustainable under current conditions but caution is advised against assuming high biodiversity interpreted as high species richness is an indicator of higher resilience. My results suggest that species rich environments could be more vulnerable due to their increased susceptibility to genetic drift, though high levels of gene flow observed could temporarily help offset genetic diversity loss.

These findings further urge consideration of genetic metrics as part of biodiversity hotspots conservation, especially as traditionally species rich regions are amongst the areas most affected by climate change (Ramírez *et al.* 2017) that according to my results harbor high genetic

diversity. Increased vulnerability can be expected as future declines across populations and species lead to loss of genetically unique populations, beneficial alleles that could be contained within them and their adaptive potential. On the other hand, increased genetic similarity in these regions may facilitate species' perseverance due to populations persisting through increased gene flow. Conservation attempts through population restocking may also benefit from naturally high rates of gene flow. Increased gene flow between populations could be advantageous to re-stocking depleted populations with individuals from other locations at a wider geographical scale, neighboring populations also benefiting from re-stocking programs.

I expect human-caused declines in genetic diversity to increase vulnerability to future environmental change (Barrett & Schluter 2008; Raffard *et al.* 2019). As variation in environmental conditions was found to be connected to genetic diversity, it is important to consider potential future implications. While this study accounts for contemporary environments and does not directly include a metric of climate change, future changes in oceanographic conditions have the potential to reorganize existing patterns of genetic diversity and likely scenarios are considered based on current results.

Low oxygen levels were found to harbor higher allelic richness, cautioning potential for loss of communities with high genetic diversity. As oxygen levels drop below survivability thresholds, these highly diverse populations may have less time to adapt as they are closer to their survivability threshold. Populations were found to be more genetically homogenous in low oxygen environments. Therefore, increased gene flow between populations could offset declines at least temporarily.

Highly productive environments were found to be more genetically similar and harbor higher genetic diversity. With low productivity zones predicted to expand under climate change

(Polovina *et al.* 2008), genetically diverse high productivity environments could be under increased threat of declines and fragmentation.

High genetic diversity regions associated with higher temperature marine regions could face rapid losses as already warm waters are pushed beyond their thresholds. Higher temperatures driving increased genetic similarity could encourage at least temporary persistence, actions such as re-stocking with thermally resistant individuals potentially having wider success.

Adaptation to environmental conditions under climate change could benefit from higher gene flow, increasing the likelihood of spreading beneficial alleles between populations. Interestingly, this could be the case under conditions of warming oceans and ocean deoxygenation and could benefit regions with higher species richness. Conservation could be more widely beneficial under high gene flow, as healthy populations can contribute to depleted populations' recovery. Under ocean acidification conditions and increased climate-change driven instances of freshening (Boyer *et al.* 2005), reorganization of pre-existing genetic diversity patterns is not expected to be linked to either salinity or pH levels directly, but genetic diversity loss is still expected indirectly due to subsequent population declines.

Conclusion

As the implications of human intervention are visible at the genetic level, I expect future adaptation to be challenging in already degraded marine environments. Adaptation is more efficient when starting from high levels of standing genetic variation as adapting from already existing, potentially already selected for, genetic information is easier than evolving new potentially beneficial mutations (Barrett & Schulter 2008). My results connect declines in

genetic diversity to human activity, implying that human behavior is inhibiting adaptive potential in marine fishes.

Widely applicable global results are a useful tool when addressing policy makers and a global audience. Decision makers' backgrounds are often from a variety of different fields, often not scientific, and these results are an easy to comprehend way to draw the attention of a global audience to issues regarding marine ecosystems. Studies at the local level will be valuable for conservation policies such as spatial planning and protected area design. Marine protected areas, an often-contentious issue when it comes to their design and implementation, neither provide adequate coverage of genetic diversity metrics (Paz-Vinas et al. 2023) nor capture genetic diversity patterns (Figuerola-Ferrando *et al.* 2023, Schmidt *et al.* 2024) at the global scale. Planning for protected areas should conserve high and protect low genetic diversity areas while forming a network of high connectivity that will contribute to genetic rescue aiming to enhance adaptive capacity across global oceans. Knowledge of both local and global parameters will help tailor conservation planning to account for both global trends and regional differences. Studies on the influence of seasonality and incorporating a metric of climate change are future research directions that could also be beneficial for adaptive management.

As pressure on marine resources continues to grow, globalization and the extractive nature of fisheries make regulatory mechanisms hard to enforce (Berkes *et al.* 2006). A global approach to marine fish conservation has become increasingly urgent. Although positive at the local level, my results caution against too much optimism about stock recoveries (Fernandes & Cook 2013; Van Gemert & Andersen 2018). These stocks tend to be managed, and their recovery process does unfortunately not extend to solving ocean-wide declines and may create a false sense of security in public opinion. A high census population size does not necessarily

indicate a viable effective population size (Frankham 1995; Palstra & Ruzzante 2008). Threatened populations could go unobserved until it is too late; thus, genetic metrics should be included in stock assessments and fishing quotas. With concerns about fisheries induced evolution and its long-term effects on recovery (Enberg *et al.* 2009), we must rethink how we monitor wild and commercially exploited species and address discrepancies in assessment methods (e.g., Schmidt *et al.* 2023). Efforts to develop ocean observation frameworks integrating genetic diversity monitoring that apply to policy should be further encouraged (e.g., Hoban *et al.* 2022; Muller-Karger *et al.* 2018).

Wide-scale patterns of genetic diversity across oceans, species and their drivers were identified despite the large amount of noise expected when analyzing such large datasets. This further demonstrates the value of open data in conservation research, the importance of encouraging open data practices and the benefits of a macrogenetic approach. My study has provided a basis for understanding processes shaping genetic diversity across global oceans at human interference and oceanographic levels. As human activity causes declines in marine fishes globally, rethinking design and development of urban coastal communities, regulating human activity in coastal waters and the high seas and transition to sustainability are increasingly urgent. Understanding spatial distribution of global genetic diversity and the relationship between genetic diversity and environment is a starting point for including genetic diversity into marine spatial planning.

Acknowledgements

I thank all the authors who made their data publicly available.

References

Alberti, M. (2015). Eco-evolutionary dynamics in an urbanizing planet. Trends Ecol. Evol., 30, 114–126.

Alberti, M., Correa, C., Marzluff, J.M., Hendry, A.P., Palkovacs, E.P., Gotanda, K.M., et al. (2017). Global urban signatures of phenotypic change in animal and plant populations. Proc. Natl. Acad. Sci., 114, 8951–8956.

Assis, J., Tyberghein, L., Bosch, S., Verbruggen, H., Serrão, E.A. & De Clerck, O. (2018). Bio-ORACLE v2.0: Extending marine data layers for bioclimatic modelling. Glob. Ecol. Biogeogr., 27, 277–284.

Barrett, R. & Schluter, D. (2008). Adaptation from standing genetic variation. Trends Ecol. Evol., 23, 38–44.

Bay, R.A. & Palumbi, S.R. (2014). Multilocus adaptation associated with heat resistance in reef-building corals. Curr. Biol., 24, 2952–2956.

Berg, P.R., Jentoft, S., Star, B., Ring, K.H., Knutsen, H., Lien, S., et al. (2015). Adaptation to low salinity promotes genomic divergence in Atlantic cod (*Gadus morhua* L.). Genome Biol. Evol., 7, 1644–1663.

Berkes, F., Hughes, T.P., Steneck, R.S., Wilson, J.A., Bellwood, D.R., Crona, B., et al. (2006). Globalization, roving bandits, and marine resources. Science, 311, 1557–1558.

Bitter, M.C., Kapsenberg, L., Gattuso, J.-P. & Pfister, C.A. (2019). Standing genetic variation fuels rapid adaptation to ocean acidification. Nat. Commun., 10, 5821.

Bivand, R.S. & Wong, D.W.S. (2018). Comparing implementations of global and local indicators of spatial association. TEST, 27, 716–748.

Blanchet, F.G., Legendre, P. & Borcard, D. (2008). Forward selection of explanatory variables. Ecology, 89, 2623–2632.

Bowen, B.W., Gaither, M.R., DiBattista, J.D., Iacchei, M., Andrews, K.R., Grant, W.S., et al. (2016). Comparative phylogeography of the ocean planet. Proc. Natl. Acad. Sci., 113, 7962–7969.

Boyer, T.P., Levitus, S., Antonov, J.I., Locarnini, R.A. & Garcia, H.E. (2005). Linear trends in salinity for the World Ocean, 1955–1998. Geophys. Res. Lett., 32, 2004GL021791.

Bürkner, P.-C. (2017). brms : An R package for Bayesian multilevel models using Stan. J. Stat. Softw., 80.

Center for International Earth Science Information Network (CIESIN) Columbia University. (2018a). Gridded population of the world, Version 4 (GPWv4): land and water area, Revision 11.

Center for International Earth Science Information Network (CIESIN) Columbia University. (2018b). Gridded population of the world, Version 4 (GPWv4): population density, Revision 11.

Chakraborty, R., Smouse, P.E. & Neel, J.V. (1988). Population amalgamation and genetic variation: Observations on artificially agglomerated tribal populations of Central and South America. Am. J. Hum. Genet., 43, 709–725.

Charlesworth, B. & Charlesworth, D. (2010). Elements of evolutionary genetics. Roberts and Company, Greenwood Village (Colo.).

Cinner, J.E., Graham, N.A.J., Huchery, C. & Macneil, M.A. (2013). Global effects of local human population density and distance to markets on the condition of coral reef fisheries. Conserv. Biol., 27, 453–458.

Clark, R. & Pinsky, M. (2024). Global patterns of genetic diversity in marine fishes. Ecol. Evol., 14, e11365

Consuegra, S., Verspoor, E., Knox, D. & García De Leániz, C. (2005). Asymmetric gene flow and the evolutionary maintenance of genetic diversity in small, peripheral Atlantic salmon populations. Conserv. Genet., 6, 823–842.

Currie, D.J. (1991). Energy and large-scale patterns of animal- and plants species richness. Am. Nat., 137, 27–49.

Dalongeville, A., Nielsen, E.S., Teske, P.R. & Von Der Heyden, S. (2022). Comparative phylogeography in a marine biodiversity hotspot provides novel insights into evolutionary processes across the Atlantic-Indian Ocean transition. Divers. Distrib., 28, 2622–2636.

De Souza Júnior, M.B., Ferreira, F.F. & De Oliveira, V.M. (2014). Effects of the spatial heterogeneity on the diversity of ecosystems with resource competition. Phys. Stat. Mech. Its Appl., 393, 312–319.

Deutsch, C., Penn, J.L. & Lucey, N. (2024). Climate, oxygen, and the future of marine biodiversity. Annu. Rev. Mar. Sci., 16, 217–245.

Deutsch, C., Penn, J.L. & Seibel, B. (2020). Metabolic trait diversity shapes marine biogeography. Nature, 585, 557–562.

Do, C., Waples, R.S., Peel, D., Macbeth, G.M., Tillett, B.J. & Ovenden, J.R. (2014). NeEstimator v2: Re-implementation of software for the estimation of contemporary effective population size (Ne) from genetic data. Mol. Ecol. Resour., 14, 209–214.

Donihue, C.M. & Lambert, M.R. (2015). Adaptive evolution in urban ecosystems. AMBIO, 44, 194–203.

Dray, S., Blanchet, G., Borcard, D., Guenard, G., Jombart, T., Larocque, G., et al. (2017). adespatial: Multivariate multiscale spatial analysis. Doi: 10.32614/CRAN.package.adespatial

Dray, S., Legendre, P. & Peres-Neto, P.R. (2006). Spatial modelling: A comprehensive framework for principal coordinate analysis of neighbour matrices (PCNM). Ecol. Model., 196, 483–493.

Duarte, C.M., Agusti, S., Barbier, E., Britten, G.L., Castilla, J.C., Gattuso, J.-P., et al. (2020). Rebuilding marine life. Nature, 580, 39–51.

Enberg, K., Jørgensen, C., Dunlop, E.S., Heino, M. & Dieckmann, U. (2009). Implications of fisheries-induced evolution for stock rebuilding and recovery. Evol. Appl., 2, 394–414.

Fernandes, P.G. & Cook, R.M. (2013). Reversal of fish stock decline in the northeast Atlantic. Curr. Biol., 23, 1432–1437.

Figuerola-Ferrando, L., Barreiro, A., Montero-Serra, I., Pagès-Escolà, M., Garrabou, J., Linares, C., et al. (2023). Global patterns and drivers of genetic diversity among marine habitat-forming species. Glob. Ecol. Biogeogr., 32, 1218–1229.

Fine, P.V.A. (2015). Ecological and evolutionary drivers of geographic variation in species diversity. Annu. Rev. Ecol. Evol. Syst., 46, 369–392.

Frankham, R. (1995). Effective population size/adult population size ratios in wildlife: A review. Genet. Res., 66, 95–107.

Gaggiotti, O.E., Bekkevold, D., Jørgensen, H.B.H., Foll, M., Carvalho, G.R., Andre, C., et al. (2009). Disentangling the effects of evolutionary, demographic, and environmental factors influencing genetic structure of natural populations: Atlantic herring as a case study. Evolution, 63, 2939–2951.

Gagnaire, P.-A., Normandeau, E., Côté, C., Møller Hansen, M. & Bernatchez, L. (2012). The genetic consequences of spatially varying selection in the panmictic American eel (*Anguilla rostrata*). Genetics, 190, 725–736.

Goudet, J. & Jombart, T. (2015). hierfstat: Estimation and tests of hierarchical F-statistics.

Halpern, B.S., Frazier, M., Potapenko, J., Casey, K.S., Koenig, K., Longo, C., et al. (2015). Spatial and temporal changes in cumulative human impacts on the world's ocean. Nat. Commun., 6, 7615.

Hare, M.P., Nunney, L., Schwartz, M.K., Ruzzante, D.E., Burford, M., Waples, R.S., et al. (2011). Understanding and estimating effective population size for practical application in marine species management: Applying effective population size estimates to marine species management. Conserv. Biol., 25, 438–449.

Harniman, R., Merritt, T.J.S., Chapman, L.J., Lesbarrères, D. & Martinez, M.L. (2013). Population differentiation of the African cyprinid *Barbus neumayeri* across dissolved oxygen regimes. Ecol. Evol., 3, 1495–1506.

Harvey, B.P., McKeown, N.J., Rastrick, S.P.S., Bertolini, C., Foggo, A., Graham, H., et al. (2016). Individual and population-level responses to ocean acidification. Sci. Rep., 6, 20194.

Hewitt, G. (2000). The genetic legacy of the Quaternary ice ages. Nature, 405, 907–913.

Hoban, S., Archer, F.I., Bertola, L.D., Bragg, J.G., Breed, M.F., Bruford, M.W., et al. (2022). Global genetic diversity status and trends: Towards a suite of Essential Biodiversity Variables (EBVs) for genetic composition. Biol. Rev., 97, 1511–1538.

Hoffmann, A., Griffin, P., Dillon, S., Catullo, R., Rane, R., Byrne, M., et al. (2015). A framework for incorporating evolutionary genomics into biodiversity conservation and management. Clim. Change Responses, 2, 1.

Jenkins, C.N. & Van Houtan, K.S. (2016). Global and regional priorities for marine biodiversity protection. Biol. Conserv., 204, 333–339.

Johnson, M.T.J. & Munshi-South, J. (2017). Evolution of life in urban environments. Science, 358, eaam8327.

Jombart, T. (2008). adegenet : A R package for the multivariate analysis of genetic markers. Bioinformatics, 24, 1403–1405.

Jombart, T. & Ahmed, I. (2011). adegenet 1.3-1: New tools for the analysis of genome-wide SNP data. Bioinformatics, 27, 3070–3071.

Jones, M.B., Slaughter, P., Nahf, R., Boettiger, C., Jones, C., Read, J., et al. (2016). dataone: R interface to the DataONE network of data repositories. Doi:10.5063/F1M61H5X

Kahilainen, A., Puurtinen M., Kotiaho J.S. (2014). Conservation implications of species–genetic diversity correlations. GECCO, 2, 315–323.

Karachaliou, E., Schmidt, C., De Greef, E., Docker, M.F. & Garroway, C.J. (2025). Urbanization is associated with reduced genetic diversity in marine fish populations. Mol. Ecol., e17711.

Kassambara A. (2023). ggcorrplot: Visualization of a correlation matrix using ’ggplot2’.R package. Doi: 10.32614/CRAN.package.ggcorrplot

Kirby, M.X. (2004). Fishing down the coast: Historical expansion and collapse of oyster fisheries along continental margins. Proc. Natl. Acad. Sci., 101, 13096–13099.

Kroodsma, D.A., Mayorga, J., Hochberg, T., Miller, N.A., Boerder, K., Ferretti, F., et al. (2018). Tracking the global footprint of fisheries. Science, 359, 904–908.

Leberg, P.L. (2002). Estimating allelic richness: Effects of sample size and bottlenecks. Mol. Ecol., 11, 2445–2449.

Li, H. (2003). Management of coastal mega-cities—a new challenge in the 21st century. Mar. Policy, 27, 333–337.

Lüdecke, D., Ben-Shachar, M., Patil, I., Waggoner, P. & Makowski, D. (2021). performance: An R package for assessment, comparison and testing of statistical models. J. Open Source Softw., 6, 3139.

Luikart, G., Sherwin, W.B., Steele, B.M. & Allendorf, F.W. (1998). Usefulness of molecular markers for detecting population bottlenecks via monitoring genetic change. Mol. Ecol., 7, 963–974.

Makowski, D., Ben-Shachar, M. & Lüdecke, D. (2019). bayestestR: Describing effects and their uncertainty, existence and significance within the Bayesian framework. J. Open Source Softw., 4, 1541.

Manel, S., Guerin, P.-E., Mouillot, D., Blanchet, S., Velez, L., Albouy, C., et al. (2020). Global determinants of freshwater and marine fish genetic diversity. Nat. Commun., 11, 692.

Mccusker, M.R. & Bentzen, P. (2010). Positive relationships between genetic diversity and abundance in fishes. Mol. Ecol., 19, 4852–4862.

McElreath, R. (2018). Statistical rethinking: A Bayesian course with examples in R and Stan. 1st edn. Chapman and Hall/CRC.

Mittell, E.A., Nakagawa, S. & Hadfield, J.D. (2015). Are molecular markers useful predictors of adaptive potential? Ecol. Lett., 18, 772–778.

Muller-Karger, F.E., Miloslavich, P., Bax, N.J., Simmons, S., Costello, M.J., Sousa Pinto, I., et al. (2018). Advancing marine biological observations and data requirements of the complementary Essential Ocean Variables (EOVs) and Essential Biodiversity Variables (EBVs) Frameworks. Front. Mar. Sci., 5, 211.

Nei, M. (1973). Analysis of gene diversity in subdivided populations. Proc. Natl. Acad. Sci., 70, 3321–3323.

Nei, M., Maruyama, T. & Chakraborty, R. (1975). The bottleneck effect and genetic variability in populations. Evolution, 29, 1–10.

Palstra, F.P. & Ruzzante, D.E. (2008). Genetic estimates of contemporary effective population size: What can they tell us about the importance of genetic stochasticity for wild population persistence? Mol. Ecol., 17, 3428–3447.

Pante, E. & Simon-Bouhet, B. (2013). marmap: A package for importing, plotting and analyzing bathymetric and topographic data in R. PLoS ONE, 8, e73051.

Paz-Vinas, I., Vandergast, A., Schmidt, C., Leigh, D., Blanchet, S., Clark, R., et al. (2023). Uneven genetic data limits biodiversity assessments in protected areas globally. EcoEvoRxiv https://doi.org/10.32942/X2ZC84

Pespeni, M.H., Sanford, E., Gaylord, B., Hill, T.M., Hosfelt, J.D., Jaris, H.K., et al. (2013). Evolutionary change during experimental ocean acidification. Proc. Natl. Acad. Sci., 110, 6937–6942.

Pinsky, M.L. & Palumbi, S.R. (2014). Meta-analysis reveals lower genetic diversity in overfished populations. Mol. Ecol., 23, 29–39.

Polovina, J.J., Howell, E.A. & Abecassis, M. (2008). Ocean’s least productive waters are expanding. Geophys. Res. Lett., 35, 2007GL031745.

QGIS Development Team. (2021). QGIS geographic information system. Open Source Geospatial Foundation Project. http://qgis.osgeo.org

R Core Team (2020). R: A language and environment for statistical computing. https://www.r-project.org/

R Core Team (2021). R: A language and environment for statistical computing. https://www.r-project.org/

Raffard, A., Santoul, F., Cucherousset, J. & Blanchet, S. (2019). The community and ecosystem consequences of intraspecific diversity: a meta-analysis. Biol. Rev., 94, 648–661.

Ramírez, F., Afán, I., Davis, L.S. & Chiaradia, A. (2017). Climate impacts on global hot spots of marine biodiversity. Sci. Adv., 3, e1601198.

Riccioni, G., Stagioni, M., Landi, M., Ferrara, G., Barbujani, G. & Tinti, F. (2013). Genetic structure of bluefin tuna in the Mediterranean Sea correlates with environmental variables. PLoS ONE, 8, e80105.

Rogers, A.D. (2000). The role of the oceanic oxygen minima in generating biodiversity in the deep sea. Deep Sea Res. Part II Top. Stud. Oceanogr., 47, 119–148.

Saenz-Agudelo, P., Dibattista, J.D., Piatek, M.J., Gaither, M.R., Harrison, H.B., Nanninga, G.B., et al. (2015). Seascape genetics along environmental gradients in the Arabian Peninsula: Insights from ddRAD sequencing of anemonefishes. Mol. Ecol., 24, 6241–6255.

Sala, E. & Knowlton, N. (2006). Global marine biodiversity trends. Annu. Rev. Environ. Resour., 31, 93–122.

Schmidt, C. & Garroway, C.J. (2021a). The conservation utility of mitochondrial genetic diversity in macrogenetic research. Conserv. Genet., 22, 323–327.

Schmidt, C. & Garroway, C.J. (2021b). The population genetics of urban and rural amphibians in North America. Mol. Ecol., 30, 3918–3929.

Schmidt, C. & Garroway, C.J. (2022). Systemic racism alters wildlife genetic diversity. Proc. Natl. Acad. Sci., 119, e2102860119.

Schmidt, C., Domaratzki, M., Kinnunen, R.P., Bowman, J. & Garroway, C.J. (2020). Continent-wide effects of urbanization on bird and mammal genetic diversity. Proc. R. Soc. B Biol. Sci., 287, 20192497.

Schmidt, C., Dray, S. & Garroway, C.J. (2022a). Genetic and species-level biodiversity patterns are linked by demography and ecological opportunity. Evolution, 76, 86–100.

Schmidt, C., Munshi-South, J., Dray, S. & Garroway, C.J. (2022b). Determinants of genetic diversity and species richness of North American amphibians. J. Biogeogr., 49, 2005–2015.

Schmidt, C., Hoban, S., Hunter, M., Paz-Vinas, I. & Garroway, C.J. (2023). Genetic diversity and IUCN Red List status. Conserv. Biol., 37, e14064.

Schmidt, C., Karachaliou, E., Vandergast, A., Crandall, E., Falgout, J., Hunter, M., et al. (2024). The global protected area network does not harbor genetically diverse populations. EcoEvoRxiv https://doi.org/10.32942/X26G7Z

Sperling, E.A., Boag, T.H., Duncan, M.I., Endriga, C.R., Marquez, J.A., Mills, D.B., et al. (2022). Breathless through time: Oxygen and animals across earth's history. Biol. Bull., 243, 184–206.

Storch, D., Bohdalková, E. & Okie, J. (2018). The more-individuals hypothesis revisited: The role of community abundance in species richness regulation and the productivity–diversity relationship. Ecol. Lett., 21, 920–937.

Sumaila, U.R., Lam, V.W.Y., Miller, D.D., Teh, L., Watson, R.A., Zeller, D., et al. (2015). Winners and losers in a world where the high seas is closed to fishing. Sci. Rep., 5, 8481.

Therkildsen, N.O., Hemmer-Hansen, J., Als, T.D., Swain, D.P., Morgan, M.J., Trippel, E.A., et al. (2013). Microevolution in time and space: SNP analysis of historical DNA reveals dynamic signatures of selection in Atlantic cod. Mol. Ecol., 22, 2424–2440.

Todd, P.A., Heery, E.C., Loke, L.H.L., Thurstan, R.H., Kotze, D.J. & Swan, C. (2019). Towards an urban marine ecology: Characterizing the drivers, patterns and processes of marine ecosystems in coastal cities. Oikos, 128, 1215–1242.

Tyberghein, L., Verbruggen, H., Pauly, K., Troupin, C., Mineur, F. & De Clerck, O. (2012). Bio-ORACLE: A global environmental dataset for marine species distribution modelling. Glob. Ecol. Biogeogr., 21, 272–281.

Van Gemert, R. & Andersen, K.H. (2018). Challenges to fisheries advice and management due to stock recovery. ICES J. Mar. Sci., 75, 1864–1870.

Weir, B.S. & Goudet, J. (2017). A unified characterization of population structure and relatedness. Genetics, 206, 2085–2103.

Worm, B., Barbier, E.B., Beaumont, N., Duffy, J.E., Folke, C., Halpern, B.S., et al. (2006). Impacts of biodiversity loss on ocean ecosystem services. Science, 314, 787–790.

Worm, B., Hilborn, R., Baum, J.K., Branch, T.A., Collie, J.S., Costello, C., et al. (2009). Rebuilding global fisheries. Science, 325, 578–585.

Worm B. & Tittensor D.P. (2018a). Drivers and predictors of global biodiversity. In: A theory of global biodiversity. Princeton University Press, Princeton, Oxford, pp. 56–92.

Worm B. & Tittensor D.P. (2018b). Observed patterns of global biodiversity. In: A theory of global biodiversity. Princeton University Press, Princeton, Oxford, pp. 10–55.

Yeager, L.A., Marchand, P., Gill, D.A., Baum, J.K. & McPherson, J.M. (2017). Marine socio-environmental covariates: Queryable global layers of environmental and anthropogenic variables for marine ecosystem studies. Ecology, 98, 1976–1976.

Table 2.1. Model summaries for human-caused and environmental associations with genetic diversity and differentiation. All model parameters had a normal prior centered on zero with a standard error of 1. Model diagnostics are presented for models at most significant km buffer per candidate driver. Diagnostics for models across all km buffers can be found in Supplemental Information 1.7. Effect sizes for each predictor are given with 95% credible intervals. The widely applicable information criterion (WAIC) and efficient approximate leave-one-out cross-validation (LOO) are an indicator of relative model fit, higher values representing better model fit. Probability of direction (pd) is a measure of the probability an effect is positive or negative. Conditional R2 is the variance explained by both fixed and random effects while marginal R2 is the variance explained only by the fixed effects. Bayesian R2 is the variance explained by both fixed and random effects divided by the total variance in the data.

Model	l-95% CI	u-95% CI	pd	Conditional R2	Marginal R2	Bayesian R2	LOOIC	WAIC
Gene Diversity ~ Primary Productivity	-0.02	0.15	93.17%	0.827	0.004	0.8271529	1255	1248.1
Gene Diversity ~ Distance to Market	-0.28	0.03	95.35%	0.829	0.016	0.8288427	1262.4	1251
Gene Diversity ~ Fishing Pressure	-0.57	-0.04	99.25%	0.333	0.052	0.3354829	27020.4	75285.7
Gene Diversity ~ Dissolved Oxygen	-0.47	0.19	80.05%	0.862	0.019	0.8622618	1021.3	1014
Gene Diversity ~ pH	-0.18	0.27	67.30%	0.842	0.007	0.8414373	1177.5	1169.3
Gene Diversity ~ Salinity	-0.23	0.79	83.85%	0.865	0.058	0.8646578	1010.1	1002
Gene Diversity ~ Temperature	-0.21	0.32	63.52%	0.858	0.009	0.8575629	1053	1047.1
Gene Diversity ~ Land Area	0.00	0.11	97.89%	0.392	0.003	0.3944736	9519.1	9608.2
Gene Diversity ~ Urbanization	-0.17	-0.01	98.80%	0.831	0.009	0.8312344	1179.7	1173.8

Gene Diversity ~ Human Impacts	**-0.11**	**0.03**	**88.05%**	**0.807**	**0.002**	**0.8065582**	**1379.1**	**1374.1**
Gene Diversity ~ Species Richness	**-0.33**	**0.22**	**68.49%**	**NA**	**NA**	**NA**	**NA**	**NA**
Allelic Richness ~ Primary Productivity	**-0.01**	**0.16**	**95.97%**	**0.755**	**0.005**	**0.7550021**	**1635.4**	**1632.8**
Allelic Richness ~ Distance to Market	**-0.15**	**0.06**	**73.76%**	**0.747**	**0.002**	**0.7472214**	**1671**	**1668.3**
Allelic Richness ~ Fishing Pressure	**-0.05**	**0.11**	**75.76%**	**NA**	**NA**	**NA**	**NA**	**NA**
Allelic Richness ~ Dissolved Oxygen	**-0.51**	**0.03**	**95.55%**	**0.782**	**0.043**	**0.7817697**	**1506**	**1506.4**
Allelic Richness ~ pH	**-0.22**	**0.28**	**62.18%**	**0.792**	**0.008**	**0.7915057**	**1471.1**	**1471.1**
Allelic Richness ~ Salinity	**-0.16**	**0.68**	**87.48%**	**0.772**	**0.059**	**0.7722535**	**1567.7**	**1565.4**
Allelic Richness ~ Temperature	**-0.06**	**0.38**	**92.40%**	**0.779**	**0.023**	**0.7790546**	**1519.3**	**1520.2**
Allelic Richness ~ Land Area	**0.02**	**0.13**	**99.39%**	**0.33**	**0.006**	**0.3350284**	**7113.5**	**6741**
Allelic Richness ~ Urbanization	**-0.15**	**0.02**	**94.23%**	**0.751**	**0.004**	**0.7502231**	**1589.8**	**1587.1**
Allelic Richness ~ Human Impacts	**-0.13**	**0.00**	**97.31%**	**0.747**	**0.004**	**0.7464375**	**1663.1**	**1661.1**
Allelic Richness ~ Species Richness	**-0.08**	**0.36**	**90.22%**	**0.763**	**0.021**	**0.7628287**	**1600**	**1593.3**
F_{ST} ~ Primary Productivity	**-0.21**	**0.03**	**91.57%**	**0.576**	**0.007**	**0.5753536**	**2209.4**	**2207**
F_{ST} ~ Distance to Market	**-0.06**	**0.31**	**90.45%**	**0.545**	**0.015**	**0.5447602**	**2305**	**2300.4**
F_{ST} ~ Fishing Pressure	**-0.06**	**0.10**	**72.44%**	**0.188**	**0.001**	**0.1985703**	**6608.2**	**4711.6**
F_{ST} ~ Dissolved Oxygen	**-0.02**	**0.87**	**97.00%**	**0.576**	**0.125**	**0.575366**	**2226.4**	**2223.3**
F_{ST} ~ pH	**-0.40**	**0.14**	**84.05%**	**0.515**	**0.018**	**0.5144288**	**2377**	**2373.8**
F_{ST} ~ Salinity	**-1.07**	**0.30**	**86.61%**	**0.594**	**0.123**	**0.5936833**	**2188.1**	**2183.5**
F_{ST} ~ Temperature	**-0.64**	**-0.01**	**97.90%**	**0.589**	**0.086**	**0.5887324**	**2189.4**	**2188.2**
F_{ST} ~ Land Area	**-0.41**	**-0.04**	**99.34%**	**0.392**	**0.04**	**0.3919392**	**24560.9**	**29105.4**
F_{ST} ~ Urbanization	**-0.12**	**0.21**	**70.50%**	**0.549**	**0.004**	**0.5488618**	**2203.0**	**2198.8**
F_{ST} ~ Human Impacts	**-0.08**	**0.12**	**65.75%**	**0.490**	**0.001**	**0.488848**	**2410.5**	**2408.7**
F_{ST} ~ Species Richness	**-0.30**	**-0.03**	**99.33%**	**0.489**	**0.024**	**0.4885272**	**2394.9**	**2392.9**
Ne ~ Primary Productivity	**-0.11**	**0.23**	**80.05%**	**0.356**	**0.006**	**0.3554855**	**1710.7**	**1704.9**

Ne ~ Distance to Market	**-0.13**	**0.12**	**53.70%**	**0.331**	**0.002**	**0.3300793**	**1738**	**1732**
Ne ~ Fishing Pressure	**-0.10**	**0.30**	**83.12%**	**0.275**	**0.01**	**0.2760769**	**5303.3**	**3646.8**
Ne ~ Dissolved Oxygen	**-0.17**	**0.35**	**76.98%**	**0.357**	**0.014**	**0.3564425**	**1710.2**	**1706.5**
Ne ~ pH	**-0.22**	**0.17**	**65.77%**	**0.337**	**0.005**	**0.3364788**	**1727.9**	**1723.8**
Ne ~ Salinity	**-0.28**	**0.23**	**64.08%**	**0.342**	**0.006**	**0.3416718**	**1723.9**	**1719.9**
Ne ~ Temperature	**-0.35**	**0.13**	**83.93%**	**0.355**	**0.017**	**0.3540018**	**1713.1**	**1709.1**
Ne ~ Land Area	**-0.15**	**0.05**	**85.12%**	**NA**	**NA**	**NA**	**NA**	**NA**
Ne ~ Urbanization	**-0.23**	**0.03**	**93.60%**	**0.275**	**0.007**	**0.2808449**	**2575.5**	**2412.4**
Ne ~ Human Impacts	**-0.05**	**0.17**	**82.38%**	**0.325**	**0.003**	**0.3239981**	**1737.1**	**1733.7**
Ne ~ Species Richness	**-0.24**	**0.04**	**92.73%**	**0.335**	**0.01**	**0.334092**	**1720.3**	**1716.8**

Table 2.2. Table of relationships between genetic diversity measures and their candidate environmental and human-caused drivers. Negative relationships noted as (-), positive relationships noted as (+) and absence of relationship noted as (None).

	Gene Diversity	**Allelic Richness**	**F_{ST}**	**Ne**
Primary Productivity	+	+	-	**None**
Distance to Market	-	**None**	+	**None**
Fishing Pressure	-	**None**	**None**	**None**
Dissolved Oxygen	**None**	-	+	**None**
pH	**None**	**None**	**None**	**None**
Salinity	**None**	**None**	**None**	**None**
Temperature	**None**	+	-	**None**
Land Area	+	+	-	**None**
Urbanization	-	-	**None**	-
Human Impacts	**None**	-	**None**	**None**

Species Richness	**None**	+	-	-

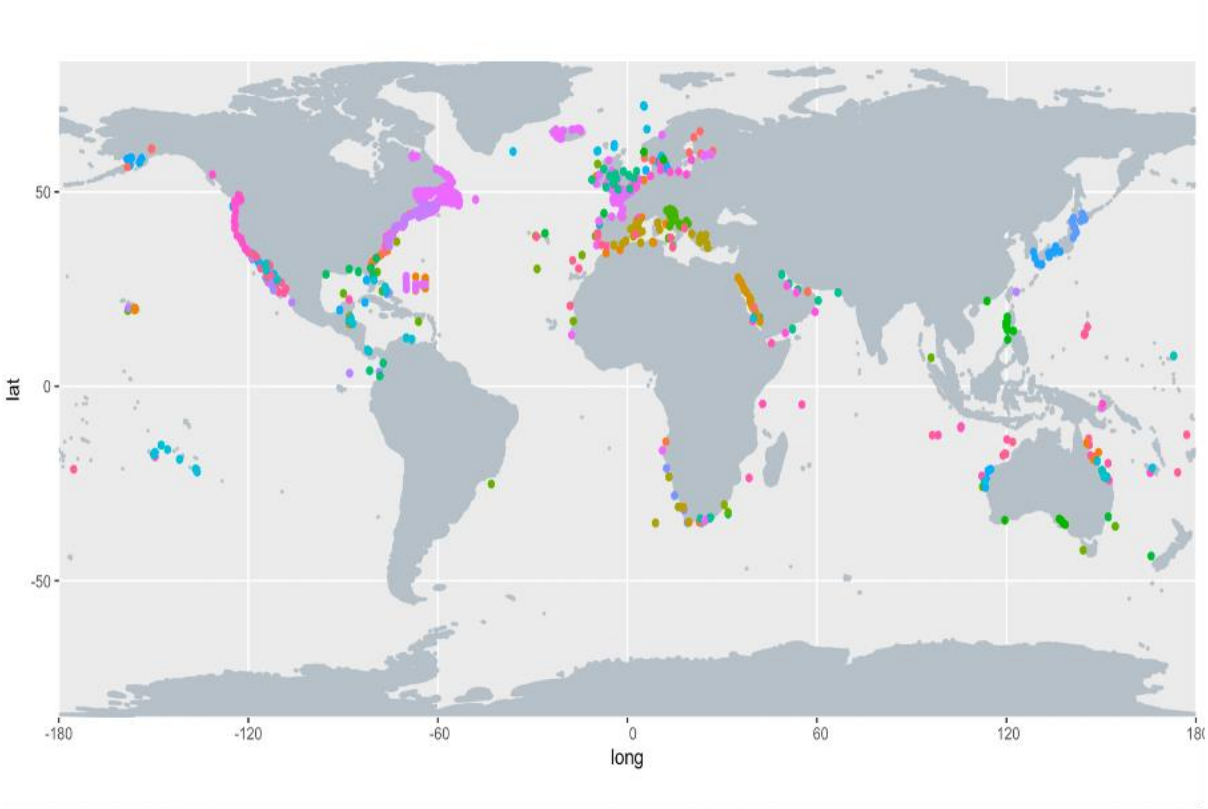

Figure 2.1. Global microsatellite dataset: 75,496 individuals, 1,143 populations and 73 fish species collected, filtered & georeferenced. Summary and reference information for data can be found in Supplemental Information 1.2. A map showing which species are represented by which color can be found in Supplemental Information 1.3. Each point represents a distinct population, and different colors represent different species.

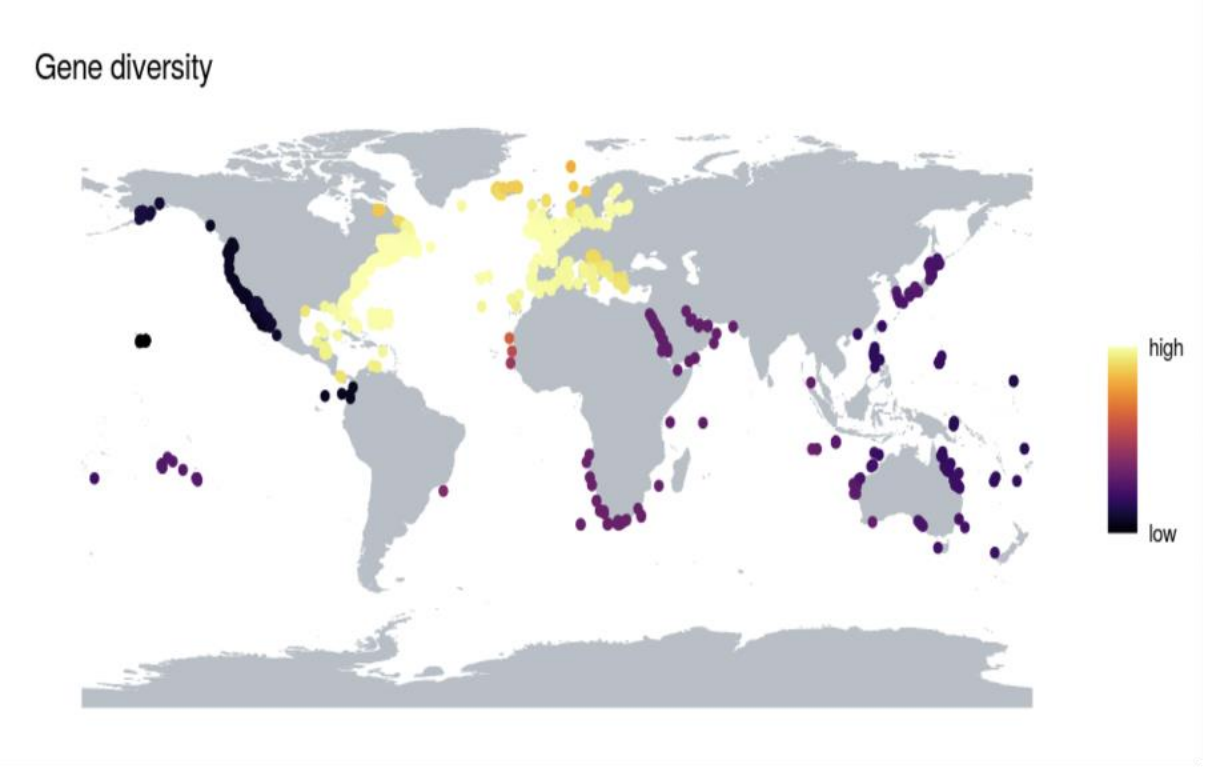

Figure 2.2. Patterns of spatial variation in gene diversity (expected heterozygosity). High levels of gene diversity found in the North Atlantic, Baltic, Mediterranean Sea, Gulf of Mexico and Caribbean regions. Intermediate values found in the South Atlantic, Indian Ocean, West and Central South Pacific Oceans and the Red Sea, with a decline along a west to east gradient. Low levels found in the East and Central North Pacific Ocean.

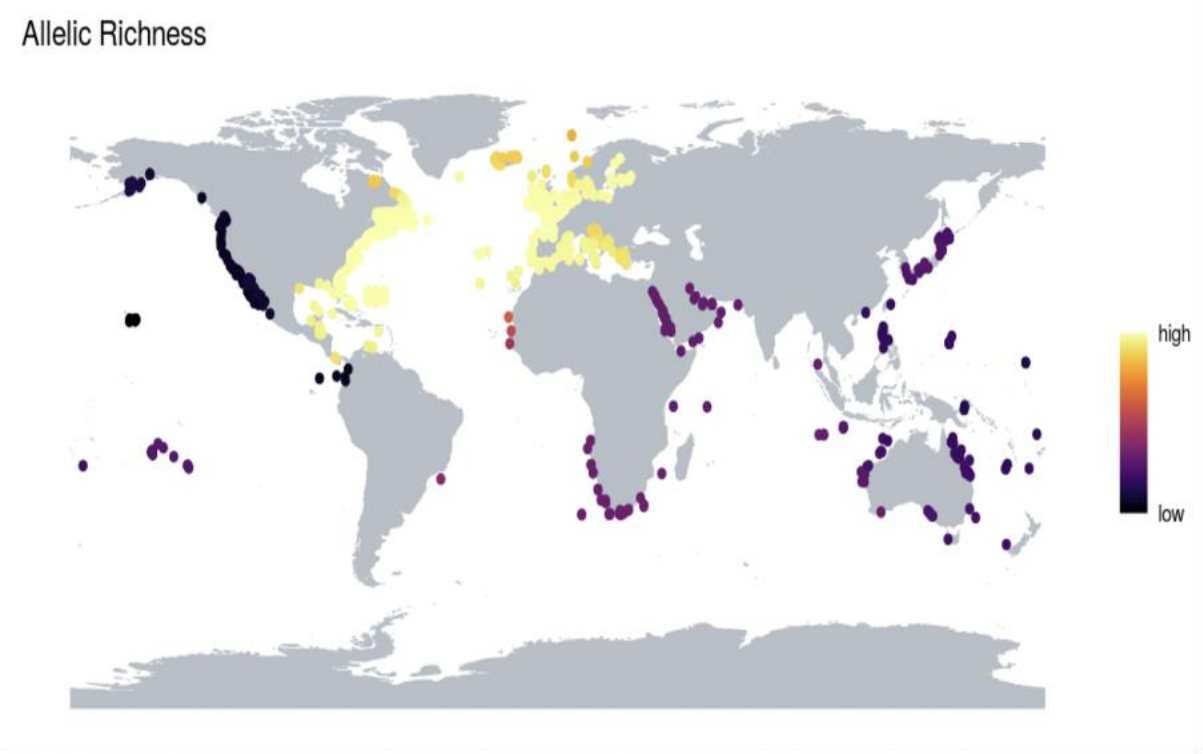

Figure 2.3. Patterns of spatial variation in allelic richness. High levels of allelic richness found in the North Atlantic, Baltic, Mediterranean Sea, Gulf of Mexico and Caribbean regions. Intermediate values found in the South Atlantic, Indian Ocean, West and Central South Pacific Oceans and the Red Sea, with a decline along a west to east gradient. Low levels found in the East and Central North Pacific Ocean.

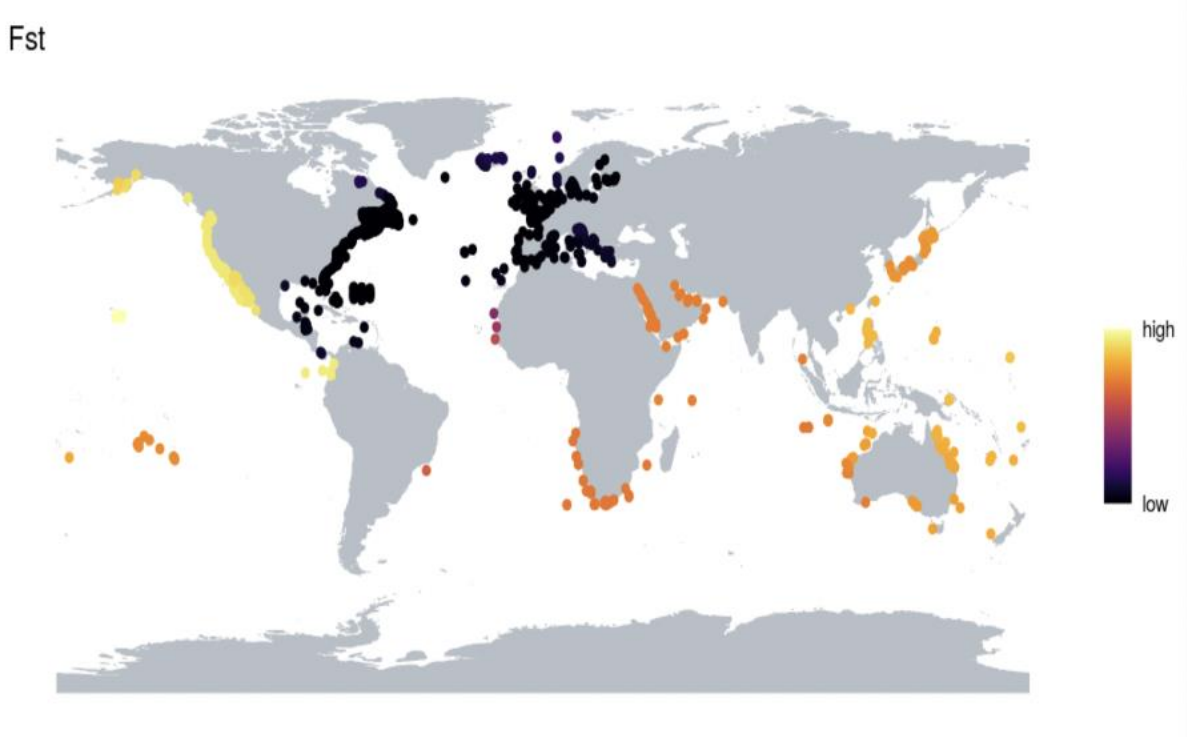

Figure 2.4. Patterns of spatial variation in F_{ST}. Low levels of F_{ST} found in the North Atlantic, Baltic, Mediterranean Sea, Gulf of Mexico and Caribbean regions. Intermediate values found in the South Atlantic, Indian Ocean, West and Central South Pacific Oceans and the Red Sea, increasing along a west to east gradient. High levels found in the East and Central North Pacific Ocean.

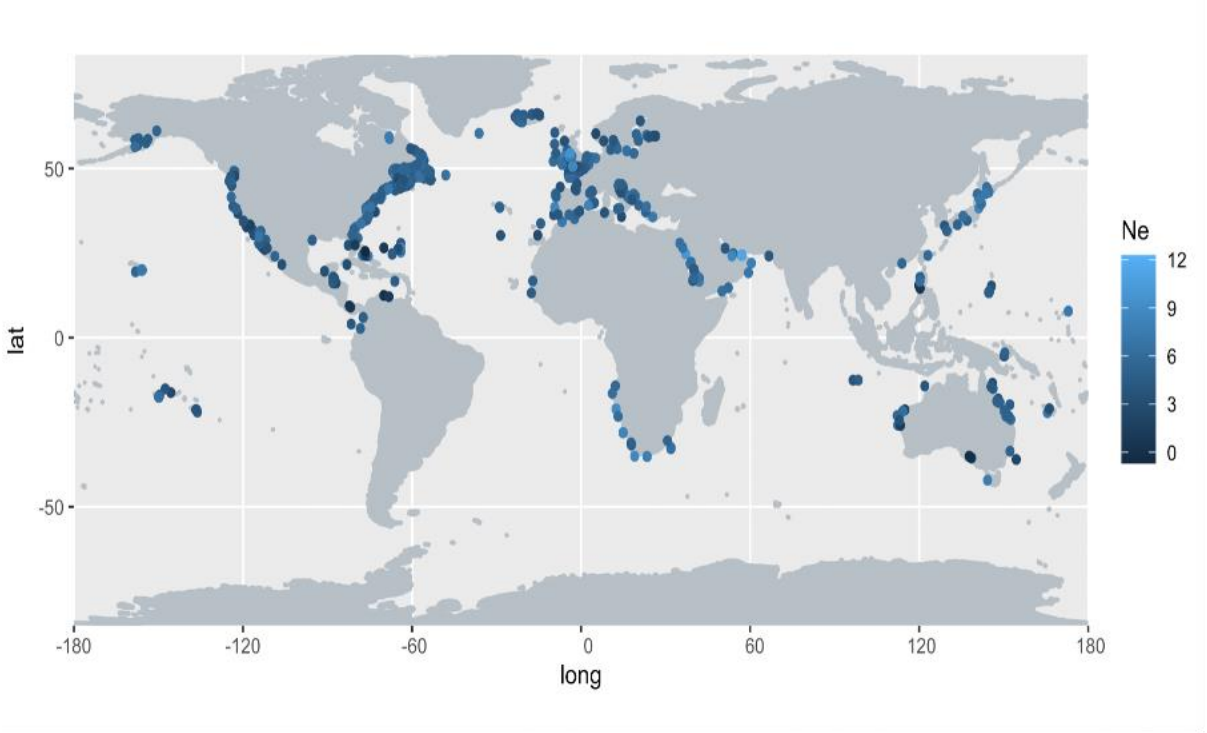

Figure 2.5. Plot of raw Ne data, no global spatial patterns detected.

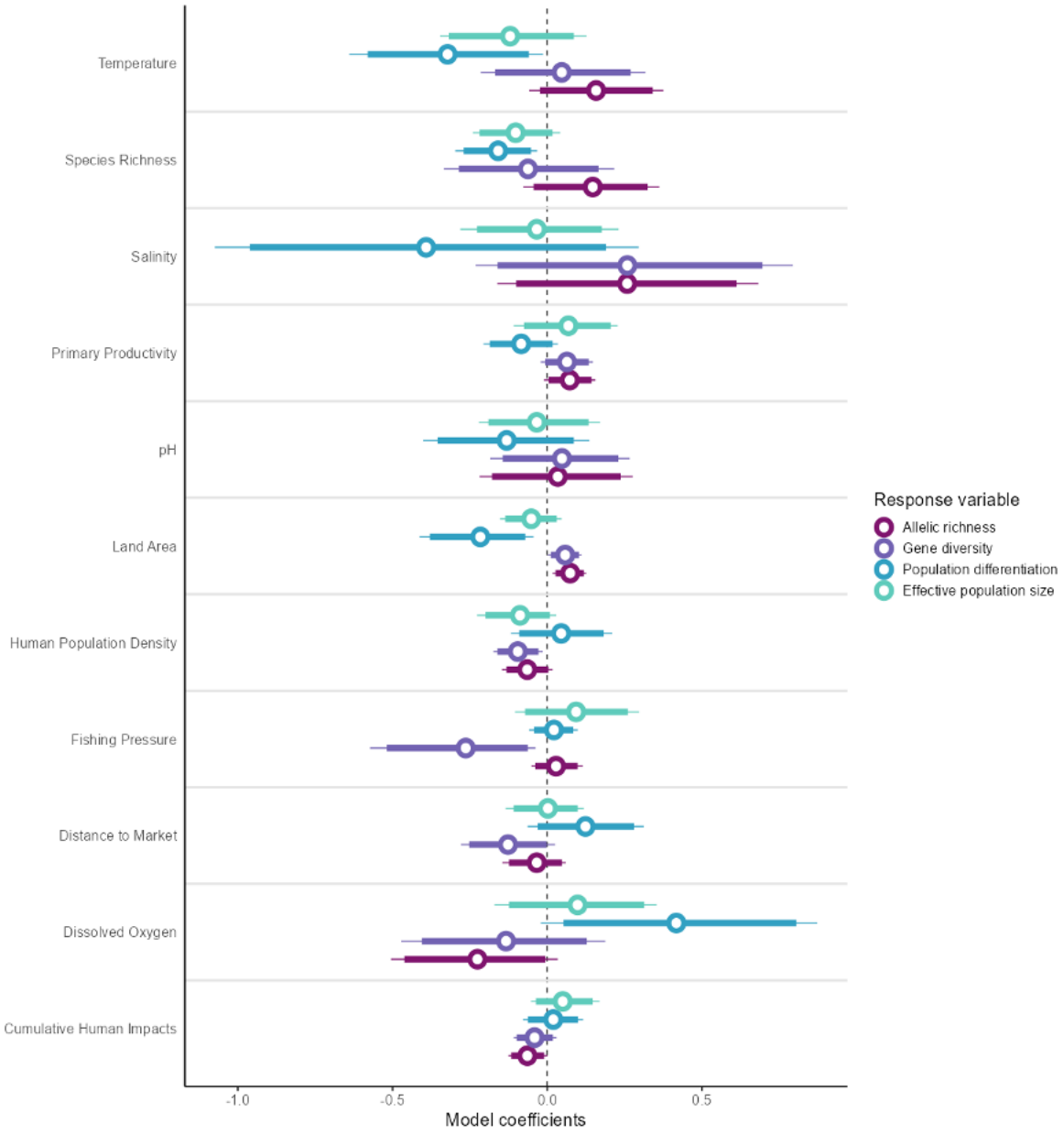

Figure 2.6. Plot of model coefficients; x axis showing candidate drivers of allelic richness, gene diversity, population differentiation (F_{ST}) and effective population size (Ne), y axis showing direction of effects. Slopes and intercepts were allowed to vary with species. Open circles are overall parameter estimates, bold lines are 90% credible intervals, and narrow lines are 95% credible intervals. Model at most significant km buffer shown.

CHAPTER 3: RECONSTRUCTING GENOME-WIDE POPULATION STRUCTURE, DEMOGRAPHY AND EVOLUTIONARY HISTORY FOR ANADROMOUS SEA LAMPREY (*PETROMYZON MARINUS*)

Karachaliou E. [1], Suurväli J.[1], de Greef E. [1], Garroway C.J. [1], Docker M.F. [1]

[1] Department of Biological Sciences, University of Manitoba, 50 Sifton Rd, Winnipeg, MB, R3T 2N2, Canada

Abstract

Anadromous sea lamprey (*Petromyzon marinus*) is of conservation concern in its native range on both coasts of the Atlantic Ocean and the Mediterranean Sea. I use genomic data to uncover population structure, status, re-construct recent and deep demographic history and inform conservation efforts. For the purpose of this study, whole genomes from 155 individuals were re-sequenced from locations throughout the sea lamprey native range. Population structure analysis confirmed trans-Atlantic divergence in anadromous sea lamprey and uncovered a distinct population in the Mediterranean Sea, encouraging designating distinct management units. Estimates of contemporary effective population size showed that vulnerability levels vary across different locations, encouraging a more localized approach to conservation. I found that all sea lamprey populations have undergone steep declines in the recent past, and a slight recent recovery in Europe, potentially indicating effectiveness of recent conservation efforts. Deep demographic history seems to be connected to Quaternary glaciation cycles.

Introduction

Native anadromous sea lamprey *Petromyzon marinus* are of conservation concern despite landlocked invasive populations in the Laurentian Great Lakes thriving (Wilkie *et al.* 2022). Due to the complexity of an anadromous life cycle, they are vulnerable to human-caused pressures, such as physical barriers affecting migration and habitat degradation (Maitland *et al.* 2015). Meanwhile commercial fisheries targeting sea lamprey operate in Europe (Almeida *et al.* 2021). These pressures vary geographically. While sea lamprey are classified as Least Concern based on their IUCN Red List assessment (Ford 2024), they experience varying threat levels across their native range (NatureServe 2024, OSPAR status assessment; BDC 2022/sea lamprey *Petromyzon marinus*), encouraging consideration of locality-specific management practices. Population genetics can provide information to assist management and conservation planning, with conservation plans directly benefiting from incorporating a genetic perspective (Dahle *et al.* 2018; Hare *et al.* 2011; Neely *et al.* 2021; Palsbøll *et al.* 2007; Xuereb *et al.* 2019). This chapter aims to produce information on sea lamprey status and demography to assist conservation efforts.

This ancient jawless vertebrate, widely distributed across diverse environmental conditions, is native to both sides of the North Atlantic Ocean. Their range extends from Labrador to the Gulf of Mexico in the west, and from the Barents Sea to Norway and Iceland, along the coast of Europe to North Africa, including the western Mediterranean and Adriatic Sea in the east (Potter *et al.* 2015). Adult sea lamprey migrate to rivers where they spawn and die, with larvae living in the sediment as filter feeders for 4-8 years until they undergo metamorphosis, after which the juvenile lamprey then migrate out to sea where they parasitize host fish species (Dawson *et al.* 2015; Hardisty & Potter 1971; Renaud & Cochran 2019). Sea

lamprey do not home to their natal streams (Bergstedt & Seelye 1995), despite natal homing as part of migration being common in other anadromous fishes (Waldman *et al.* 2008). The ecological importance of sea lamprey has been reviewed in Docker *et al.* (2015); during their burrowing larval stage and nest building spawning phase, lampreys resuspend sediment and contribute to nutrient cycling in freshwater systems while simultaneously functioning as ecosystem engineers. Lampreys then die after spawning, releasing marine-derived nutrients into nutrient poor freshwater ecosystems. Sea lamprey are a high nutritional value food source for other animals, including humans, reflected by their cultural importance in Europe (reviewed in Docker *et al.* 2015).

Anadromous sea lamprey form two genetically distinct populations, each located on the east (European) and west (North American) coasts of the Atlantic Ocean, with the West Atlantic population expressing higher genetic diversity (Almada *et al.* 2008; Bryan *et al.* 2005; Genner *et al.* 2012; Rodriguez-Munõz *et al.* 2004) and effective population size (Almada *et al*. 2008; Bryan *et al.* 2005; Genner *et al.* 2012). Estimates of recent evolutionary history based on mitochondrial DNA (mtDNA) show an increase in genetic diversity ~125,000 years ago in the East Atlantic and ~500,000 years ago in the West Atlantic (Genner *et al.* 2012). To date, no significant genetic structure has been found within either Atlantic coast (Almada *et al.* 2008; Bryan *et al.* 2005; Genner *et al.* 2012; Rodriguez-Munõz *et al.* 2004, Waldman *et al.* 2006). However, these studies have been based on nuclear microsatellite markers and mitochondrial DNA haplotypes. While informative, these markers have their limitations. Common issues encountered include homoplasy, maternal inheritance of mtDNA, and variation in mutation rates and null alleles in microsatellites confounding analysis (Morin *et al.* 2004). In comparison single nucleotide polymorphisms (SNPs) can provide better and more uniform genome-wide coverage and higher

quality data (Morin *et al.* 2004). Compared to microsatellites SNPs can better estimate population genetic structure and diversity (Zimmerman *et al.* 2020). The first study using whole genome sequencing focused on sea lamprey in the East Atlantic and found them to be genetically homogenous while displaying signs of latitudinal isolation by distance (Baltazar-Soares *et al.* 2023). Isolation by distance has also been observed in anadromous Pacific lamprey *Entosphenus tridentatus* (Hess *et al.* 2013; Lin *et al.* 2008; Spice *et al.* 2012), a species with a similar life history. Limited dispersal at sea could prevent total panmixia as it is thought that sea lamprey rarely disperse far from the continental shelf and slope, though individuals are sometimes found in deeper waters (Mateus *et al.* 2021). Different feeding preferences could cause differences in dispersal distances between these species despite their similar life histories (Mateus *et al.* 2021; Spice *et al.* 2012); longer attachment of blood feeding sea lamprey to hosts could increase distance traveled at sea compared to flesh feeding Pacific lamprey that change hosts more often (Potter & Hilliard 1987; Spice *et al.* 2012). Thus, I expect sea lamprey to be more genetically homogenous than Pacific lamprey. Additionally, non-genetic evidence of finer-scale population structure in sea lamprey has been found in Portugal (Lança *et al.* 2014), with differences in morphology and fatty acid signatures linked to local seabed conditions suggesting limited movement during their marine phase (Lança *et al.* 2014).

Secondary contact has yet to be detected in sea lamprey, though transoceanic movement has been recorded in Pacific lamprey (Murauskas *et al.* 2019). Genner *et al.* (2012) suggested that either colonization of Europe by North American migrants took place during the last 150,000 years, or that divergence between the two populations occurred earlier in their evolutionary history and an environmentally driven bottleneck occurred in Europe ~125,000 years ago. This study was based on mtDNA markers and its extension by using the vast amount

of information contained in genome-wide data would be beneficial. Complex trans-Atlantic genetic structure is known for diadromous species such as Atlantic salmon *Salmo salar* (Lehnert *et al.* 2020). Adding sea lamprey evolutionary history to existing literature will enable further understanding of the evolutionary trajectory followed by diadromous species, movement across the North Atlantic and potential drivers of evolution.

Next-generation sequencing will provide an opportunity to further our current understanding of the genetic structure underlying anadromous sea lamprey oceanographically diverse habitat range and life history. My study uses medium coverage whole-genome sequencing data, low coverage data already successfully used in studies of demography and population structure (e.g., Foote *et al.* 2016). Genome-wide data has also been recommended for use in effective population size (Ne) estimates in marine fish species (Hemmer-Hansen *et al.* 2014). Interpreted as the rate at which a population is susceptible to genetic drift, Ne in marine fishes can be very different from census population size (Hauser & Carvalho 2008) and is a valuable metric of population vulnerability to loss of genetic diversity that can inform conservation and policy (Hoban *et al.* 2021). To date, studies that account for anadromous sea lamprey range-wide have focused on mitochondrial and microsatellite markers, which reflect genetic diversity less well than SNPs (Morin *et al.* 2004; Zimmerman *et al.* 2020). Therefore, this study aims to use genomic data by scanning across the genome for SNPs, for individuals from across the entire sea lamprey native distribution range. Data produced by this study on demography, population structure and status will hopefully contribute to management and conservation efforts. Adding data on sea lamprey to existing literature will enable further understanding of evolutionary trajectory of diadromous species in the North Atlantic Ocean.

Information on fine-scale sea lamprey population structure, status and demographic history will also help identify management units and plan for a more localized approach to native sea lamprey conservation. At the local level, questions have arisen regarding the origin of sea lamprey in Icelandic waters, especially due to their increased occurrences in the northernmost parts of their range on both coasts (Pereira *et al.* 201; Van Leeuwen *et al.* 2021). Population substructure could also be driven by unique life histories. Europe lacks any known landlocked sea lamprey populations, though feeding in fresh water before outmigration to sea has been recorded (Docker & Potter 2019). While sea lamprey in Sweden inhabit a transitional zone, they are not established in the lower salinity waters of the Baltic Sea (Mikael Svensson, SLU Swedish Species Information Centre, pers. comm.). I will therefore delineate sea lamprey population structure. The primary driver of genetic divergence is predicted to be trans-Atlantic distance, with the secondary driver being distance between the Atlantic and the Mediterranean basin. Population genetic structure within and between each Atlantic coast is predicted to be linked to limits to dispersal and oceanographic barriers. Specialized life history characteristics will likely drive further structure. Following the detection of genetically distinct populations, I will estimate contemporary effective population size. This information can then be used to address local conservation concerns such as those regarding the status of Swedish sea lamprey (Mikael Svensson, SLU Swedish Species Information Centre, pers. comm.). I will then reconstruct demographic history in the recent past. I expect recent demographic history to be linked to habitat degradation and barriers to migration, with the East Atlantic reflecting the effects of existing lamprey fisheries in Europe. I will then reconstruct demographic history in the distant past and estimate the timing of the historical split between East and West Atlantic sea lamprey populations. I expect the split between East and West Atlantic coasts and deep demographic

trajectory to be connected to geographic isolation, climatic oscillations and glaciation events, results reflecting their timing.

Methods

Samples

Anadromous sea lamprey samples were acquired from collaborators across their native range. My final dataset consists of 155 individuals, 109 from the East Atlantic, 40 from the West Atlantic and 6 from the Mediterranean Sea. Although most individuals were collected as adults, larval samples were collected in Germany, Maryland and New Brunswick. Detailed sample information is included in Table 3.1.

DNA extraction and sequencing

High quality DNA was extracted using the DNeasy Blood and Tissue Kit (Qiagen). Samples were sent for PCR-free library construction and whole genome sequencing to The Centre for Applied Genomics (TCAG; The Hospital for Sick Children, Toronto, Canada), where sequencing was conducted on two Illumina NovaSeq 6000 runs utilizing paired 150-bp reads. Sequencing was undertaken as part of a larger scale study including both anadromous sea lamprey across their native range and freshwater resident sea lamprey in the Great Lakes.

Data processing, variant calling and filtering

Raw sequence data had already been demultiplexed and converted to files in .fastq format upon delivery. Subsequent data processing, SNP variant calling and filtering were done by Jaanus Suurväli (University of Manitoba) as part of the larger scale study. The program fastp (v0.23.1) (Chen *et al.* 2018) was used for adapter removal and quality trimming. Reads shorter than 70bp after trimming were removed. Reports from FastQC (v0.11.9) (Andrew 2015) and

fastp itself were used for quality control. Samples sequenced on multiple lanes were merged into two files per individual, one containing forward reads and one containing reverse reads. Reads were aligned to the sea lamprey germline assembly kPetMar1.pri (Timoshevskaya *et al.* 2023) using BWA mem (v0.7.17) (Li & Durbin 2009) under default settings. MarkDuplicates from Picard (v2.26.3) (Broad Institute 2019) was used to mark and remove duplicate reads, while AddOrReplaceReadGroups was used to add a unique read group identifier to each sample. Duplicate reads, reads that could not be mapped as a pair, those that failed platform/vendor quality checks or had a quality score below 20 whilst mapping were removed using samtools (v1.15.1) (Danecek *et al.* 2021). Platypus (v0.8.1) (WGS500 Consortium *et al.* 2014) was then used to identify genetic variants. Computation was split across parallel nodes. The resulting .vcf files were merged using vcf-concat from vcftools (v0.1.16) (Danecek *et al.* 2011). Only the variants passing all default filters suggested by Platypus were used for downstream analyses. As Platypus often reports haplotypes rather than individual SNPs, bcftools norm was used to decompose these into one variant per row.

Sample names in the file were changed to names more appropriate for downstream analyses using bcftools (v1.11) (Danecek *et al.* 2021). The subsequent file was sorted using bcftools sort. Anadromous sea lamprey individuals used for this thesis were extracted from the full sea lamprey file and reordered based on location using bcftools view from bcftools v0.1.11. Genetic sites with no polymorphisms present after removal of the samples not used in this study were excluded using SelectVariants from the Genome Analysis Toolkit (GATK; v4.2.6.1) (McKenna *et al.* 2010). Using bcftools annotate, database identifiers for all chromosomes and chromosome-associated genome fragments were replaced with their human-readable RefSeq counterparts, such as “chr01”. Any sites with data missing from more than 20% of all individuals

were filtered out with bcftools. Fourteen genome fragments unassigned to specific chromosomes were left with either no variants or with only a single variant, and were excluded from further analysis. The following analysis is based on the anadromous only data file and undertaken in R version 4.3.2 (R Core Team 2023) unless otherwise mentioned.

Population structure

For population structure analysis, anadromous sea lamprey data were filtered using bcftools v0.1.16. Only biallelic SNPs were retained and the effects of linkage disequilibrium and low frequency alleles were controlled for by removing all sites with minor allele frequency < 5%, and thinning data to retain only variants separated by at least 1000 nucleotides (1 kb). Since my dataset includes larval samples, I tested for relatedness between individuals by calculating identity by descent for the West Atlantic and East Atlantic/Mediterranean separately in the R package dartR (v2.9.7) (Mijangos *et al.* 2022). Using Vcftools v0.1.16, nucleotide diversity and pairwise F_{ST} as per Weir & Cockerham (1984) were estimated on a windowed basis (window size 20kb). These metrics were calculated for within and between East, West Atlantic coasts and the Mediterranean Sea, and all distinct locations. Principal component analysis (PCA) was run using the glpca() command from the R package adegenet (v2.1.10) (Jombart 2008; Jombart & Ahmed 2011). The snmf function in LEA (Frichot & François 2015), a method that can estimate ancestry coefficients, was used to create plots of population genetic structure based on candidate K=2-6 ancestral populations. PCAs and snmf analysis in LEA were run on the entire dataset, an East Atlantic/Mediterranean and a West Atlantic subset.

Two larval samples from Maryland, PR2021_02 and PR2021_04, were found to be outliers to the rest of North American sea lamprey individuals. Therefore, species identification was confirmed using genetic data. The BLAST tool (https://blast.ncbi.nlm.nih.gov/Blast.cgi) was

used to search the NCBI nucleotide database for sequences similar to mitochondrial DNA reads extracted from the two outliers.

Effective population size, demographic & evolutionary history

Contemporary effective population size (Ne) was calculated using the ldNe command from the R package strataG (Archer *et al.* 2017). Ne is estimated based on linkage disequilibrium across pairs of unlinked presumed neutral loci using Pearson's squared correlation coefficient as a metric (Waples *et al.* 2016). Effective population size can be interpreted as the number of breeding individuals in an ideal population losing genetic variation at the same rate as the observed population and the susceptibility of a population to the influence of genetic drift. Scaffolds smaller than 100,000 bp were removed using a custom shell script to control for noise created by small scaffolds. Non-biallelic SNPs, non-polymorphic sites and sites where more than 10% of individuals have missing data were removed using bcftools v0.1.11. Five distinct .vcf files each containing a 10,000 SNP subsample were produced using plink v1.90b as a thinning tool, due to limits on the number of SNPs StrataG can process. These files were used as input for strataG to calculate contemporary Ne across five distinct runs, for each distinct location.

Recent demographic history was uncovered by estimating recent trajectory of Ne using GONE (Santiago *et al*. 2020) for the most genetically distinct populations identified. Analysis was run on the 50 longest scaffolds due to limitations on the number of scaffolds that GONE can use as input and problems caused by large differences in number of SNPs between scaffolds. Data was subset to include only the longest 50 scaffolds using bcftools. Non biallelic SNPs, non-polymorphic sites and sites where more than 10% of individuals have missing data were also removed using bcftools. Files were converted to .ped and .map format using plink as input data. GONE estimates were calculated across 100 replicate runs using Haldane's correction for genetic

distance and a maximum of 50,000 SNPs analyzed per chromosome. Recombination rate was set to 6.39 cM per Mb based on dividing the combined genetic distance of 5570.25 cM calculated for sea lamprey whole embryos by Smith & Keinath (2015) with 80% of the length of the kPetMar1.pri reference germline genome, as this corresponds to the expected approximate size of the somatic genome.

Past sea lamprey demographic trajectory, population expansions, declines and the split between the East and West Atlantic coasts were estimated using SMC++ (Terhorst *et al.* 2017). Rates of coalescence are used to infer effective population size (Ne) across time, with time to the most recent common ancestor estimated across segments of the genome while taking into account linkage disequilibrium. Scaffolds smaller than 100,000 bp were removed using a custom shell script. VCF files were converted to .smc format using the vcf2smc command. Masking excluded indels, sites with quality value below 30 and left only biallelic sites in the data. The demographic trajectory for each genetically distinct population was determined by 100 replicate model runs using the estimate command. Per generation mutation rate for sea lamprey is unknown to date, while generally SNP mutation rate varies between 10^{-8} and 10^{-9} (Brumfield *et al.* 2003). A per generation mutation rate of 10^{-9} was selected based on the spectrum consistent with current literature that used SNP data for fish species (Barth *et al*. 2017; Nikolic *et al.* 2020; Yamasaki *et al.* 2020). Further parameters were defined using the following commands set to default values --regularization-penalty 6.0 and --thinning 1000 * log(n) where n is sample size. Runs were estimated across a timescale of $10–10^{6}$ years ago. The split between the East and West Atlantic populations was estimated using 10 individuals per Atlantic coast as recommended by Terhorst *et al.* (2017). One hundred replicate model runs were generated for each Atlantic coast using the estimate command under the same parameters as above. The joint frequency spectrum

for both populations was created using the vcf2smc command. Split timing, shared and diverged evolutionary trajectory between Atlantic coasts was then estimated using the split command for 100 replicate model runs, using the 100 per population replicate model runs and their joint frequency spectrum as input. Preliminary plots were created using the plot command and based on an average generation time of 9 years (Hardisty & Potter 1971). This command also produces a file with the input data to create the plot. The data file created by the plot command was used as input to visualize and plot models in R.

Results

A total of 155 anadromous sea lamprey samples were successfully sequenced across their known distributional range. The resulting whole-genome sequencing of ~16.5x mean read depth (maximum read depth reported per sample ~ 37.4x and minimum read depth reported per sample ~8.6x) enabled the location of 11,986,442 genetic variants out of which 10,658,258 are biallelic SNPs.

Population structure

I found no related individuals in the East Atlantic/Mediterranean region. In the West Atlantic region, I found two larval individuals from Maryland to be related (PR2021_02, PR2021_04) however they are not full siblings (mean probability of identity by decent = 0.36, where duplicate samples = 1 and first-degree relatives = 0.5). It is not expected that two related individuals in a dataset this size (n = 155) will bias population structure estimates. F_{ST} estimates (Weir & Cockerham 1984) detected significant differentiation between East and West Atlantic coasts with mean F_{ST} = 0.255. Significant genetic differentiation was also found between the Mediterranean and the rest of European locations with mean F_{ST} = 0.014, although it is not as

high as differentiation between West and East Atlantic coasts. Differentiation between the West Atlantic and Mediterranean was mean F_{ST} = 0.177. No strong differentiation was detected between or within West Atlantic locations or between or within East Atlantic locations. Mean F_{ST} estimates between locations are shown in Table 3.2. and within locations in Tables 3.3., 3.4. and 3.5.

Nucleotide diversity was found to be higher in the East Atlantic (PI: 7.643e-05) compared to the West Atlantic (PI: 1.125e-04), the opposite of what was expected from the literature (Almada *et al.* 2008; Bryan *et al.* 2005; Genner *et al.* 2012; Rodriguez-Munõz *et al.* 2004). No significant differences were found within respective Atlantic coasts and detailed estimates are provided in Table 3.6.

Population structure was visually explored with admixture plots of K = 2-6 candidate ancestral populations and PCA plots that agreed with admixture results. Analysis was run separately for the entire dataset (Figure 3.1., Supplemental Information 2.1-2.3), the West Atlantic coast (Figure 3.2., Supplemental Information 2.4-2.6) and the East Atlantic coast/Mediterranean Sea region (Figure 3.3., Supplemental Information 2.7-2.9). Ancestry coefficient estimates placed East Atlantic and West Atlantic coasts as separate populations, as expected based on available literature (Almada *et al.* 2008; Bryan *et al.* 2005; Genner *et al.* 2012; Rodriguez-Munõz *et al.* 2004). No clear signs of trans-Atlantic gene flow were found, but neither can the possibility of its existence be rejected. Potential evidence of low-level gene flow could also be a signal of shared ancestry.

The Mediterranean was found to be genetically distinct from other European locations on the Atlantic coast. This distinct Mediterranean population still displays substantial gene flow with the rest of Europe. Results could also be interpreted as recent colonization of the

Mediterranean by Atlantic sea lamprey leading to high levels of common ancestry. Individuals from Lough Conn in Ireland were also found to cluster separately. Individuals at this location have been observed feeding in freshwater, contrary to typical anadromous sea lamprey behavior. Not all individuals at this location were genetically distinct, perhaps because not all sea lamprey at this location display this behavior. Ancestry analysis also found that sea lamprey from the French Atlantic coast are the most genetically unique, with the rest of East Atlantic locations being highly admixed. Sea lamprey in Sweden, thought to avoid lower salinity waters of the Baltic, and Icelandic sea lamprey are also part of the highly admixed East Atlantic population.

A latitudinal gradient in population structure was observed on the West Atlantic coast, with individuals from the most northern (Massachusetts and New Brunswick) and most southern (Maryland, North Carolina) locations clustering together. Due to substantial gene flow, they cannot be considered distinct populations and seem to be a case of isolation by distance. Two larval samples from Maryland (Little Patuxent River) are outliers to the rest of North American individuals. BLAST results matched mtDNA reads from these samples to sea lamprey sequences in the NCBI nucleotide database at the 100% level. Therefore, these samples from Maryland are not a case of misidentification at the species level during sampling efforts due to difficulties in discerning identifying features in larval individuals.

Effective population size, demographic & evolutionary history

Contemporary Ne was estimated as such: 5 runs X 10k SNP subsample X location. Estimates were then averaged across these 5 runs. A table of results across all 5 runs can be seen in Table 3.7. Gene flow is known to bias Ne estimates (Wang & Whitlock 2003). When migration rate is low the influence of genetic drift is more significant at the local level (Hare et al. 2011), so I calculated Ne per distinct location.

Sea lamprey on the French Mediterranean coast display a low Ne value of ~24, in line with being described as locally quasi extinct (Baglinière *et al.* 2020). Low Ne is also observed in Sweden (Ne ~61), Maryland (Ne ~38) and Germany (Ne ~3). High Ne values are found in Portugal (Ne ~6,997), Massachusetts (Ne ~9,758), the French Atlantic coast (Ne ~2,260) and Ireland (Ne ~1,186). Intermediate values were observed in Iceland (Ne ~358) and North Carolina (Ne ~151). Ne calculated for Iceland is likely an arbitrary number because spawning at this location has not been observed, and sea lamprey here are believed to be straying from other locations (Pereira *et al.* 2012). In New Brunswick Ne was found to be = Inf across all replicate runs. Inability to calculate Ne at this location is likely due to weak genetic drift signal causing higher error likelihood in large populations (Hare *et al.* 2011). Sea lamprey are listed as Secure in the region (S5; NatureServe 2024), so a large population size is realistic.

GONE was run on the two main genetically distinct East Atlantic (115 individuals including the Mediterranean Sea; Figure 3.4) and West Atlantic (40 individuals; Figure 3.5) populations. The Mediterranean Sea (6 individuals from the French Mediterranean coast; Figure 3.6) and France (21 individuals from the French Atlantic coast; Figure 3.7) were also considered genetically distinct enough locations to estimate their recent Ne separately. GONE reports time in generations and an average generation time for sea lamprey is 9 years (Hardisty & Potter 1971). One hundred generations (900 years) ago the West Atlantic showed a steady incline in Ne starting at ~50,000 and peaking at 125,000 around 50 generations (450 years) ago. Ne then remained around the same level following a rapid decline towards the present. Over the last 100 generations (900 years) sea lamprey in the West Atlantic initially expanded but declined rapidly in the recent past. Ne in the East Atlantic started at ~150,000 100 generations (900 years) ago. Ne then slowly declined until the occurrence of a sharp bottleneck ~80 generations (720 years)

ago. Ne remained extremely low until recovery started ~30 generations (270 years) ago raising Ne to around ~25,000. Over the last 100 generations (900 years) sea lamprey in the East Atlantic have experienced a steep decline but have shown signs of recent recovery.

Ne in France follows the same trajectory seen in the entirety of the East Atlantic, indicating that despite sea lamprey in France being found to be genetically more "unique" than the rest of sea lamprey in Europe, they share the same recent demographic history. This supports French sea lamprey not being genetically different enough to be considered a distinct population. Ne in the Mediterranean was estimated at ~5,000 around 100 generations (900 years) ago. Ne steadily increased, peaking at ~12,500 ~50 generations (450 years) ago. Ne remained fairly steady until ~20 generations (180 years) ago after which it sharply declined. Mediterranean sea lamprey initially underwent a steady expansion, followed by a substantial recent decline. Despite substantial gene flow with the rest of Europe the Mediterranean follows a different recent evolutionary trajectory, confirming that it is a distinct population. It is also noted that the Mediterranean, at least until 100 generations (900 years) ago, has been the smaller sea lamprey population.

When estimating past demographic and evolutionary history using SMC++ I initially considered the East Atlantic (115 individuals), West Atlantic (40 individuals), Mediterranean Sea (6 individuals) and France (21 individuals) as distinct populations, aiming to use all individuals per population for analysis. It is noted that SMC++ does not perform well in the near past and the time point at which I consider models to start becoming unreliable is between 100,000-10,000 years ago. In the West Atlantic (Figure 3.8., Supplemental Information 2.10) ~10,000,000 years ago Ne was estimated at ~700,000. After a slight decline, ~2,000,000 years ago Ne displayed a sharp increase to an Ne of ~50,000,000. Ne decreased abruptly to ~100,000

at ~500,000 years ago. A second increase in Ne followed, peaking ~100,000 years ago with Ne ~700,000, after which Ne is seen to steadily decrease as it gets closer to the present. Sea lamprey in the West Atlantic have undergone a rapid historical population expansion followed by an equally abrupt bottleneck. These were followed by a second, less pronounced, population expansion. The second decline in the more recent past is worth taking into consideration but less reliable as it is closer to the present. In the Mediterranean (Figure 3.9., Supplemental Information 2.11) Ne starts at ~900,000 at ~10,000,000 years ago followed by a slight decline, recovery to its initial numbers and eventual decline to a low of slightly under ~100,000 around 1,000,000 years ago. At this time an abrupt increase in Ne occurs, peaking at ~900,000 at ~800,000 years ago followed by a decline to a low of ~100,000 at ~500,000 years ago. A second rapid increase and subsequent decline in Ne to similar peak and low values is observed at ~250,000 years ago and ~100,000 years ago respectively. Ne is then observed to steadily increase as time gets closer to the present but this more recent expansion is likely unreliable. After a slight population decline and recovery, sea lamprey in the Mediterranean have undergone two successive population expansions and bottlenecks. The more recent population recovery, though unreliable, is worth taking into consideration.

Sea lamprey in France (Figure 3.9., Supplemental Information 2.12) follow the same demographic trajectory as the Mediterranean, with the more recent but unreliable population recovery being more pronounced. Due to large sample size and computational limitations, I was unable to estimate demographic history across the entirety of the East Atlantic simultaneously, so results were inferred based on estimations made using subsets. The Mediterranean and France, the two most genetically distinct locations in the East shared the same demographic trajectory with ten individuals from Portugal (Figure 3.9., Supplemental Information 2.13) (also used for

trans-Atlantic split estimates and discussed below), a representative of the highly admixed East Atlantic population. This indicates that the East Atlantic/Mediterranean region shared the same deep demographic history until ~100,000 years ago. To further support this theory of shared ancestry I estimated demographic trajectory for an additional highly admixed East Atlantic location, Germany (8 individuals; Figure 3.9, Supplemental Information 2.14). Sea lamprey in Germany share the same demographic history with the rest of the East Atlantic up until the point of the second population expansion, after which they do not undergo the second population bottleneck and signal gets lost at around ~10,000 years ago. Thus, shared demographic history in the East Atlantic can be confidently supported up to the point of initiation of this second population bottleneck. After this time point, either there has been divergence in demographic trajectory in Germany that could not be uncovered by this studies' analysis of population genetic structure, or demographic history remains shared and SMC++ was not able to computationally capture more recent evolutionary processes for this specific location.

To estimate the trans-Atlantic split, I used 10 individuals from New Brunswick (Figure 3.8, Supplemental Information 2.15) representing the West Atlantic coast and 10 individuals from the Vouga River in Portugal (Figure 3.9, Supplemental Information 2.13) representing the East Atlantic coast. When demographic history for these two subsets was considered separately prior to estimating the trans-Atlantic split, sea lamprey from New Brunswick followed the same trajectory as the entirety of the West Atlantic with an exception that the second, more recent but also unreliable, population decline was not detected as signal was lost. Sea lamprey from Portugal shared the same demographic trajectory with the Mediterranean, Germany and France, with signal starting to become unreliable at ~100,000 years ago. Split estimates (Figure 3.10) showed that in the ancestral population Ne started at ~900,000 at ~10,000,000 years ago,

followed by a decline, recovery to around its initial numbers and eventually declined to a low of ~100,000 at ~1,500,000 years ago right before the onset of the trans-Atlantic split. The split between the two Atlantic coasts occurred ~1,500,000 years ago. Post split the West Atlantic population experienced a rapid increase in Ne, estimated at ~50,000,000, followed by a rapid decline to a low of ~200,000 at ~200,000 years ago. Ne then starts to increase until signal starts to get lost ~100,000 years ago. Post split the East Atlantic population starts with a low Ne of ~10,000. Ne then increased to ~500,000 followed by a decline to ~80,000 at ~500,000 years ago. This is followed by a second increase to ~1,000,000 at ~300,000 years ago after which Ne is seen to decline again until signal is lost ~100,000 years ago. It is noted that SMC++ assumes a clean split with no subsequent gene flow. As population structure results showed little evidence of transatlantic gene flow, secondary contact is unlikely therefore this analysis method is appropriate to the study system. Still, further testing for secondary contact using other analysis methods is recommended.

Discussion

My results on population structure confirm trans-Atlantic divergence in anadromous sea lamprey and uncover a genetically distinct population in the Mediterranean Sea. Estimates of contemporary effective population size show varying levels of vulnerability across the sea lamprey native range. Reconstruction of demographic trajectory in the recent past detected steep recent declines across all sea lamprey populations. Reconstruction of deep demographic history estimated the trans-Atlantic split to have taken place ~1,500,000 years ago and that the East Atlantic seems to be the ancestral population.

Populations are thought not to be highly structured and mainly driven by isolation by distance due to limits to dispersal. Results are similar to what is known for Pacific lamprey, with sea lamprey considered more likely to mix with nearby locations due to traveling greater distances out to sea (Spice *et al.* 2012). Anadromous sea lamprey also seem to display similarities to population structure known for sea lamprey in the Great Lakes, under increased scientific scrutiny due to control programs targeting invasive populations (Marsden & Siefkes 2019). Despite their ability to disperse across long distances, landlocked sea lamprey display wide-scale structure, with Lake Ontario, Lake Erie, and the upper Great Lakes all being considered genetically distinct (reviewed in Docker *et al.* 2021). Genetically distinct populations identified in the East Atlantic, West Atlantic and Mediterranean Sea mirror this pattern of broad-scale population structure. Anadromous sea lamprey can travel long distances at sea, but dispersal seems to have its limitations (Mateus *et al.* 2021). Therefore, genetic structure should be broad-scale rather than finer-scale. However non-genetic evidence of finer-scale structure of anadromous sea lamprey has been found in Portugal, likely connected to local oceanographic conditions (Lança *et al.* 2014). Local differences in phenotypic and demographic characteristics have also been observed in landlocked sea lamprey (reviewed in Docker *et al.* 2021). Despite little overall evidence of finer-scale genetic structure in the Great Lakes (reviewed in Docker *et al.* 2021), only one study used a genomic approach and was not specifically designed to answer this question (Sard *et al.* 2020). Although my results did not detect clear finer-scale genetic structure, genetically different individuals were found in Lough Conn in Ireland and on the French Atlantic coast. It seems that sea lamprey population genetic structure is more relevant at a wider scale across both its native and invasive range.

Recent literature using next generation sequencing data on anadromous sea lamprey populations for the first time (Baltazar-Soares *et al.* 2023) reported a single admixed population in the East Atlantic showing signs of isolation by distance driven by latitude. By using an extended sampling effort my results have uncovered population structure further than what was previously known. There are sea lamprey spawning locations that were not included in my study (e.g., Italy, Spain), supporting the potential for additional undiscovered genetic structure. My results place Icelandic sea lamprey as part of the highly admixed East Atlantic population; lamprey could either be straying into Icelandic waters from nearby locations, or spawning occurs but has not yet been detected. Further research is needed to trace back the exact origin of sea lamprey found in Iceland.

Management units are important to consider for conservation and sustainable harvesting (Benestan 2019). My results add whole genome information to support previous studies that identified genetically distinct anadromous sea lamprey on the West and East Atlantic coasts, supporting their designation as distinct management units (Almada *et al.* 2008; Bryan *et al.* 2005; Genner *et al.* 2012; Mateus *et al.* 2012; Rodriguez-Munõz *et al.* 2004). Thus, management should plan to account for per coast differences in human pressures, environmental conditions, the presence of a targeted fishery and differences in demographic history.

My study also uncovered a genetically different Mediterranean population on the French coast, unfortunately after its loss (Baglinière *et al.* 2020). Oceanographic barriers in the Atlantic to Mediterranean transition zone can cause disruption to gene flow (Patarnello *et al.* 2007) and it seems to also be the case for sea lamprey. This could also be the case for historic divergence in European brook lamprey (*Lampetra planeri*); individuals found in rivers on the Mediterranean coast display lower genetic diversity and are substantially genetically different to Atlantic

locations (Rougemont *et al.* 2021). Despite being freshwater resident, they are thought to have originated from their anadromous counterparts since gone locally extinct in the Mediterranean (Potter *et al.* 2015). Anadromous *Alosa fallax* shad are also divergent between Atlantic and Mediterranean locations, thought to be a result of past glaciation events (Sabatino *et al.* 2022). Although sea lamprey sightings are rare in the Mediterranean (Çevik *et al.* 2010; Clavero *et al.* 2014; Economidis *et al.* 1999; Holčík *et al.* 2004; Karachle & Machias 2014; Mizzan & Vianello 2007; Rafrafi-Nouira *et al.* 2015; Thessalou-Legaki *et al*. 2012, reviewed in Petetta *et al.* 2019), it is thought that spawning still occurs in Italy despite their threatened status (Petetta *et al.* 2019). Conservation of Italian sea lamprey should be prioritized before they are also lost at this location. Located further east than the French Mediterranean coast, Italian sea lamprey are more likely to belong to the Mediterranean population than be genetically similar to the rest of Europe. As Mediterranean sea lamprey were found to be a naturally smaller population on the French coast, this would likely also apply to Italy. Sea lamprey in Italy are probably subject to pressures similar to those that caused extinction on the French Mediterranean coast. The Mediterranean coastline is densely populated and highly developed since antiquity, many areas also subject to overtourism (Mandić & Petrić 2021). This contributes to pollution, overfishing and unsustainable urban development in the basin, an enclosed sea with limited inflow and outflow. High levels of urbanization lead to more concentrated damming and polluting of rivers, preventing spawning. Little information is available for sea lamprey in the Mediterranean, encouraging efforts to understand more about their life history in the basin and identify existing spawning locations. It has been proposed that gene flow between East Atlantic and Mediterranean sea lamprey is facilitated by highly migratory host species crossing the Strait of Gibraltar (Petetta *et al.* 2019).

It is considered that anadromy in sea lamprey could be facultative and linked to abundance of suitable prey, and as a result, a freshwater body of water of adequate size is necessary for feeding in fresh water (Docker & Potter 2019). To date sea lamprey have only been known to have successfully become freshwater resident in the Laurentian Great Lakes, Lake Champlain, Finger Lakes and Oneida Lake (Docker & Potter 2019) and freshwater-feeding individuals in Ireland are thought to eventually migrate out to sea (King & O'Gorman 2018). Finding that some lamprey at locations with known freshwater-feeding behavior are genetically distinct implies that transition to freshwater residency could be underway. This possibility is supported by the ability of both sea lamprey and other anadromous lamprey species to establish freshwater populations and the existence of paired species, where freshwater resident species are thought to have evolved from their anadromous counterparts (Docker & Potter 2019; Potter *et al.* 2015). Further research into potential underlying adaptive mechanisms is needed to uncover why sea lamprey at this location are genetically different.

Morphological species identification of the two larval individuals from Maryland (Little Patuxent River) identified as outliers to the rest of North America as sea lamprey was further confirmed using genetic data. However, species identification at the genetic level was examined using mtDNA, a marker that is maternally inherited, so these individuals could be hybrids. The distribution range of sea lamprey overlaps with two non-parasitic lamprey species in Maryland; least brook lamprey *Lampetra aepyptera* and American brook lamprey *Lethenteron appendix* (https://www.marylandbiodiversity.com). Least brook and American brook lamprey adults are morphologically different to adult sea lamprey. Adult hybrids would therefore likely be morphologically different to the parent species, but larvae would be harder to distinguish preventing identification as hybrids in the field. While hybridization in fishes is relatively

common (Montanari *et al.* 2016; Scribner *et al.* 2000), hybrids between genera are less frequently encountered but not unheard of (Montanari *et al.* 2016). Heterospecific mating associations in co-occurring lampreys (Johnson *et al.* 2015) increase the possibility of hybrids. However, despite accounts of American brook lamprey spawning alongside sea lamprey in the Great Lakes (Morman 1979), hybridization experiments failed to produce viable larvae, reaching only the blastula stage (Piavis *et al.* 1970). Sea and American brook lamprey hybrids are not considered likely but may not be impossible. No literature or accounts of least brook and sea lamprey hybridization attempts exist to date.

Translocation of individuals can help restore declining or rebuild locally extinct populations (Novak *et al.* 2021). Restocking sea lamprey between Atlantic coasts is advised against as populations are genetically different, although to date this has not been considered. Restocking is feasible within each Atlantic coast at locations with high genetic similarity, such as closely situated locations in the West Atlantic where effects of isolation by distance are low and locations found to be highly admixed in the East Atlantic. Successful efforts in local translocation of Pacific lamprey within the Columbia River system (Hess *et al.* 2022) encourage considering this conservation method for sea lamprey. Based on my study, restocking between the Mediterranean and the rest of Europe is not advisable. Despite retaining substantial gene flow and shared deep evolutionary history, these locations are genetically different and follow divergent recent evolutionary paths. The locally extinct population on the French Mediterranean coast could be restored by restocking with individuals from Italy, though research would first be needed to determine their genetic similarity. Additionally, potential plans would depend on the recovery of sea lamprey in Italy and addressing the regional threats they face.

Restocking is currently being considered as a conservation measure to enhance low sea lamprey numbers in Sweden (Mikael Svensson, SLU Swedish Species Information Centre, pers. comm.). Based on my results this would be possible because sea lamprey in Sweden were not found to be genetically different and are part of the highly admixed East Atlantic population. Therefore, individuals could be theoretically translocated from Ireland, Germany and Portugal, though realistically German lamprey are too vulnerable for removal of individuals. Restocking with individuals from France is not recommended as they were found to be slightly genetically different. Sea lamprey from Spain were not included in this study and could be a candidate source of individuals for translocation pending further investigation.

Sea lamprey on the French Mediterranean coast have been described as quasi extinct since time of sampling in the early 2000s (Baglinière *et al.* 2020), with Ne before local extinction estimated at ~24 individuals. Swedish sea lamprey, for which there are concerns regarding their Endangered status (OSPAR status assessment; BDC 2022/sea lamprey *Petromyzon marinus*) and low numbers (Mikael Svensson, SLU Swedish Species Information Centre, pers. comm.) display a concerningly low Ne at ~61 individuals. Sea lamprey in Maryland are considered Apparently Secure (S4; NatureServe 2024) but conversely display an Ne of ~38 individuals. Although sea lamprey occurrences are frequent in the North Sea (Elliott *et al.* 2021), they are Near Threatened in Germany (OSPAR status assessment; BDC 2022/sea lamprey *Petromyzon marinus*) and very few adults have been observed migrating upstream of the River Rhine (Baer *et al.* 2018). This makes Ne ~3 found in Germany extremely worrying and could indicate imminent population collapse. Lower Ne was also found in North Carolina (Ne ~151). Sea lamprey in North Carolina are considered Vulnerable (S3; NatureServe 2024), reflecting lower Ne observed at this location. Higher Ne values are found in Portugal (Ne ~6,997), Massachusetts (Ne ~9,758), the French

Atlantic coast (Ne ~2,260) and Ireland (Ne ~1,186). High Ne values in Portugal and France are expected as they support some of the main sea lamprey populations (Mateus *et al.* 2012). Though these results are encouraging, they are commercially valuable and subject to overfishing in the region (reviewed in Maitland *et al.* 2015). Sea lamprey are considered vulnerable in Portugal and near threatened in France based on their OSPAR status assessment (BDC 2022/sea lamprey *Petromyzon marinus*). Caution is advised against overly positive interpretation of high Ne found at these European locations as sea lamprey is considered at risk in Europe mainly due to habitat loss (Mateus *et al.* 2012). Therefore, although high contemporary Ne estimates seem to indicate positive population status at these locations, current numbers are unlikely to reflect historical abundance. Sea lamprey are considered Apparently Secure in Massachusetts (S4; NatureServe 2024), though current abundance is thought to reflect around 50% of historical numbers due to habitat loss from damming and pollution (Hartel *et al.* 2002). High abundance in the Connecticut River (Kynard & Horgan 2019) reflects detected high Ne value. Sea lamprey are Near Threatened in Ireland based on their OSPAR status assessment (BDC 2022/sea lamprey *Petromyzon marinus*) and display intermediate Ne values.

Population monitoring in the wild can be logistically challenging. Ne can be a useful tool to assess the status of a population, as populations with low Ne are susceptible to inbreeding and loss of genetic diversity and evolutionary potential (Frankham *et al.* 2002), and therefore more vulnerable. Effective population size has been considered an indicator of census population size, the ratio of their relationship (Ne/Nc) approximately 0.1 (Frankham 1995; Palstra & Ruzzante 2008). However, this relationship generalizes across biologically very different species and taxa. As the Ne/Nc ratio is not known for sea lamprey, when using Ne to inform conservation efforts I interpret it as a vulnerability metric. Low Ne is an indicator of vulnerability, while high Ne is an

indicator of resilience. At the local level, Ne estimates suggest that sea lamprey in Sweden, Maryland and Germany should be prioritized for conservation and restoration efforts. Portugal and Massachusetts, locations that support large numbers of sea lamprey, are the more resilient locations based on their effective population size. This finding is encouraging, as planning for a sustainable Portuguese fishery would not be realistic if this location was vulnerable. While some locations may have historically supported naturally smaller populations than others, it is clear that human pressures have shaped current status and larger population sizes could help buffer against further declines. Both Portugal and the state of Massachusetts have the opportunity to protect these abundant sea lamprey through appropriate conservation measures.

To date, management plans specific to anadromous sea lamprey are regional. Management is implemented in Portugal and France in connection to local fisheries (Almeida *et al.* 2021), and a conservation-oriented plan exists in the Connecticut River (CRASC 2018). Commercial fisheries currently only operate on the East Atlantic coast in Portugal, Spain and France, and while there are historical accounts of harvest by immigrants of European descent on the West Atlantic coast no commercial fisheries for the species currently operate in the area (Almeida *et al.* 2021). My results support the benefits of locality specific sea lamprey conservation measures.

Over the last 100 generations (900 years) sea lamprey in the East Atlantic experienced a steep decline but seem to be recovering in more recent years. Widespread fisheries in Europe (Almeida *et al.* 2021) could have contributed to this decline, as well as increased habitat degradation and barriers to migration due to an expanding human population and infrastructure. Conservation efforts in recent years (reviewed in Almeida *et al.* 2021) could have contributed to recent recovery. I also found the East Atlantic to have a larger effective population size in the

recent past than the West, the opposite of what has been suggested to date (Almada *et al.* 2008; Bryan *et al.* 2005; Genner *et al.* 2012). Despite the East Atlantic population having declined significantly in the recent past while supporting a commercial fishery, it seems that recovery is underway and that it could benefit from a higher effective population size. These findings encourage scaling up conservation efforts. Based on my results, the East Atlantic is a distinct population with varying levels of gene flow across locations. It therefore could benefit from a unified European conservation effort aimed at supporting both a sustainable population and sustainable fisheries.

Over the last 100 generations (900 years) sea lamprey in the West Atlantic, a genetically distinct population, seem to have undergone a steady expansion followed by a steep decline in recent years. This recent decline is likely connected to habitat degradation and barriers to migration as a result of increased urban development and increased fishing pressure affecting sea lamprey prey species. The rapid recent decline has occurred despite lack of fisheries directly targeting sea lamprey (Maitland *et al.* 2015). Highly urbanized North American coastlines and river systems could be intensifying threats known for native lampreys (Maitland et al. 2015, Clemens et al. 2021, Moser et al. 2021). These threats have been extensively reviewed (Clemens *et al.* 2021; Maitland *et al.* 2015; Moser *et al.* 2021); effects of barriers to migration, pollution, habitat loss and degradation likely contributing most to the observed decline. The recent population recovery in Europe, potentially connected to conservation efforts, encourages scaling up more recent conservation efforts across the sea lamprey North American range (e.g., CRASC 2018; Hogg *et al.* 2013; Kynard & Horgan 2019; Livermore *et al.* 2017) and raising awareness on the importance of this species to help curb declines.

Over the last 100 generations (900 years), Mediterranean sea lamprey initially expanded but the population has declined rapidly in the recent past, as expected based on local extinction on the French Mediterranean coast where my samples originated from. Based on my results the Mediterranean is a distinct population that maintains high rates of gene flow with the rest of Europe. It thus should be considered separately when it comes to conservation and management. Since the Mediterranean diverged from the rest of Europe at least 100 generations (900 years) ago, it has been the smaller population compared to those situated on the West and East Atlantic coasts. This is likely due to high historical human presence in the basin and derived pressures, and the smaller enclosed Mediterranean Sea having a more limited ability to support large populations. Sea lamprey are already considered scarce in the Mediterranean and my results have now uncovered the population's historically small size and recent decline. The recent recovery detected in the rest of Europe may have benefited from management efforts implemented in connection to local fisheries, encouraging initiating conservation efforts in the Mediterranean. Gene flow between the Mediterranean and the rest of Europe may also benefit future recovery.

According to split estimates, the East Atlantic is closest to the ancestral population. Sea lamprey likely originated on the East Atlantic coast, and after an initial population bottleneck, subsequent recovery and second bottleneck, populations split due to sea lamprey colonization of the West Atlantic coast. Nucleotide diversity was found to be higher in the East Atlantic than the West Atlantic, an expected result of a founder effect, further supporting the ancestral status of the East Atlantic. Sea lamprey in the West Atlantic then likely encountered favorable environmental conditions and rapidly evolved and expanded. Migration could also have been instigated by a large number of individuals and occurred over a long period of time with the East Atlantic population continuously contributing to the West Atlantic gene pool thus preventing low

effective population size at that time. A population bottleneck followed, and then recovery in the more recent past. Sea lamprey in the East Atlantic continued to have a smaller population size post split. They then underwent two consecutive population expansions and bottlenecks.

Results place the trans-Atlantic split between East and West around 1,500,000 years ago, supporting a proposed hypothesis from the literature suggesting that based on mtDNA divergence occurred early in evolutionary history (Genner *et al.* 2012), and reflecting an overlapping time scale despite the difference in markers used. A subsequent bottleneck connected to the last glacial period in Europe ~125,000 years ago (Genner *et al.* 2012), the existence of which is also supported by Almada *et al*. (2008), is further supported by my detection of a population decline in Europe ~100,000 years ago. A second hypothesis from the literature suggesting colonization of Europe by North American migrants during the last 150,000 years (Genner *et al.* 2012) is not supported by this analysis due to differences in estimated split timing, and East Atlantic sea lamprey being placed as the ancestral population. Transatlantic divergence in anadromous sea lamprey displays a similar timing to what is known for Atlantic salmon; estimated to have occurred ~1,670,000 years ago by Rougemont & Bernatchez (2018) and > 1,000,000 years ago by Nilsson et al. (2001). Both Atlantic salmon (Rougemont & Bernatchez 2018) and sea lamprey transcontinental divergence took place during the Pleistocene, a geological period that started ~2,580,000 and lasted until ~11,000 years ago (Gibbard *et al.* 2010). Sea lamprey divergence and subsequent isolation can be placed at the initiation of the Quaternary glaciation cycles, a period thought to have shaped present day genetic diversity, population structure and evolutionary history in the Northern Hemisphere (Hewitt 2000). Timing of the historical decline detected for East Atlantic sea lamprey ~80 generations (720 years) ago corresponds to the timing of the “Little Ice Age”, a prolonged period of low temperatures and

extreme weather events linked to declines in cod fisheries in the North Atlantic (Rose *et al.* 2019). The absence of a similarly timed decline in the West Atlantic is likely due to the cooling effects of the "Little Ice Age" being amplified in the North Atlantic European region (Lehner *et al.* 2013).

Sea lamprey in the Mediterranean likely originated from the main East Atlantic population more recently, supported by results showing shared past but divergent recent demographic history. Lower Ne in the Mediterranean could indicate limits to population expansion in a smaller, enclosed sea. Divergence between Mediterranean and East Atlantic sea lamprey seems to have occurred between 900 and 100,000 years ago (based on an average generation time of nine years according to the sea lamprey life span as seen in Hardisty & Potter 1971), likely taking place during the last glacial period. This timing connects Mediterranean sea lamprey divergence to known glacial refugia in southern Europe (Hewitt 1999; Hewitt 2000), a pattern followed by other lamprey species (Mateus *et al*. 2016).

This study has described population structure, reconstructed demographic and evolutionary history, and estimated per location vulnerability to declines in genetic diversity across the anadromous sea lamprey distribution range to inform conservation. My results highlight the value of taking into account both past and present evolutionary trajectory while simultaneously considering effective population size as part of a conservation genetic approach. An ancient fish species, sea lamprey have persevered through population declines, contemporary and historical environmental changes. My results provide information that will contribute to management efforts to ensure sea lamprey survival in a changing future.

Acknowledgements

I would like to thank Ted Castro-Santos (U.S. Geological Survey), Mike Wilkie (Wilfrid Laurier University), Guillaume Evanno (INRA), Magnús Jóhannsson (MFRI), Benóný Jónsson (MFRI), Jan Baer (LAZBW), Fiona Bracken (University College Dublin), Catarina Mateus (Universidade de Évora), Thomas Evans (St. Mary's College of Maryland), Michael Fisk (NC Wildlife Resources Commission), Jeremy McCargo (NC Wildlife Resources Commission), Gabriela M. Hogue (North Carolina Museum of Natural Sciences), Mikael Svensson (SLU), Elisabeth Thysell (County Administrative Board in Halland) for the anadromous sea lamprey samples. I would like to thank Arfa Khan, Jessie L. Ogden, Phil Grayson and Matthew Thorstensen for the technical support as part of this project. This research was enabled in part by support provided by the Prairies DRI and the Digital Research Alliance of Canada (alliancecan.ca).

References

Almada, V.C., Pereira, A.M., Robalo, J.I., Fonseca, J.P., Levy, A., Maia, C., et al. (2008). Mitochondrial DNA fails to reveal genetic structure in sea-lampreys along European shores. Mol. Phylogenet. Evol., 46, 391–396.

Almeida, P.R., Arakawa, H., Aronsuu, K., Baker, C., Blair, S.-R., Beaulaton, L., et al. (2021). Lamprey fisheries: History, trends and management. J. Gt. Lakes Res., 47, S159–S185.

Andrew S. (2015). FastQC: A quality control tool for high throughput sequence data [Online]. http://www.bioinformatics.babraham.ac.uk/projects/fastqc/

Archer, F.I., Adams, P.E. & Schneiders, B.B. (2017). stratag : An R package for manipulating, summarizing and analysing population genetic data. Mol. Ecol. Resour., 17, 5–11.

Baer, J., Hartmann, F. & Brinker, A. (2018). Abiotic triggers for sea and river lamprey spawning migration and juvenile outmigration in the River Rhine, Germany. Ecol. Freshw. Fish, 27, 988–998.

Baglinière, J.L., Lasne, E, Evanno, G. & Beaulaton, L. (2020). La lamproie marine. In: Les Poissons d'eau douce de France. (coord. Keith, P., Poulet, N., Denys, G., Changeux, T., Feunteun, É. & Persat, H.). Muséum national d'Histoire naturelle, Paris ; Biotope, Mèze,, pp. 265–267.

Baltazar-Soares, M., Britton, J.R., Pinder, A., Harrison, A.J., Nunn, A.D., Quintella, B.R., et al. (2023). Seascape genomics reveals limited dispersal and suggests spatially varying selection among European populations of sea lamprey (*Petromyzon marinus*). Evol. Appl., 16, 1169–1183.

Barth, J.M.I., Damerau, M., Matschiner, M., Jentoft, S. & Hanel, R. (2017). Genomic differentiation and demographic histories of Atlantic and Indo-Pacific yellowfin tuna (*Thunnus albacares*) populations. Genome Biol. Evol., 9, 1084–1098.

Benestan, L. (2019). Population genomics applied to fishery management and conservation. In: Population genomics: Marine organisms, population genomics (eds. Oleksiak, M.F. & Rajora, O.P.). Springer International Publishing, Cham, pp. 399–421.

Bergstedt, R.A. & Seelye, J.G. (1995). Evidence for lack of homing by sea lampreys. Trans. Am. Fish. Soc., 124, 235–239.

Broad Institute. (2019). Picard toolkit. https://broadinstitute.github.io/picard/

Brumfield, R.T., Beerli, P., Nickerson, D.A. & Edwards, S.V. (2003). The utility of single nucleotide polymorphisms in inferences of population history. Trends Ecol. Evol., 18, 249–256.

Bryan, M.B., Zalinski, D., Filcek, K.B., Libants, S., Li, W. & Scribner, K.T. (2005). Patterns of invasion and colonization of the sea lamprey (*Petromyzon marinus*) in North America as revealed by microsatellite genotypes. Mol. Ecol., 14, 3757–3773.

Çevik, C., Ergüden, D. & Tekelioğlu, N. (2010). Confirmation of the presence of the sea lamprey, *Petromyzon marinus* Linnaeus, 1758 in the Levantine Sea (Petromyzoniformes: Petromyzonidae). Zool. Middle East, 49, 107–108.

Chen, S., Zhou, Y., Chen, Y. & Gu, J. (2018). fastp: An ultra-fast all-in-one FASTQ preprocessor. Bioinformatics, 34, i884–i890.

Clavero, M., Adrados, B., Calzada, J. & Jácome-Flores, M. (2014). On the presence of *Petromyzon marinus* in Oued Moulouya (Morocco). Cybium, 38, 307-308.

Clemens, B.J., Arakawa, H., Baker, C., Coghlan, S., Kucheryavyy, A., Lampman, R., et al. (2021). Management of anadromous lampreys: Common threats, different approaches. J. Gt. Lakes Res., 47, S129–S146.

CRASC (Connecticut River Atlantic Salmon Commission). (2018). Connecticut River anadromous sea lamprey management plan. CRASC Sea Lamprey Subcomm. Sunderland Mass.

Dahle, G., Johansen, T., Westgaard, J.-I., Aglen, A. & Glover, K.A. (2018). Genetic management of mixed-stock fisheries “real-time”: The case of the largest remaining cod fishery operating in the Atlantic in 2007–2017. Fish. Res., 205, 77–85.

Danecek, P., Auton, A., Abecasis, G., Albers, C.A., Banks, E., DePristo, M.A., et al. (2011). The variant call format and VCFtools. Bioinformatics, 27, 2156–2158.

Danecek, P., Bonfield, J.K., Liddle, J., Marshall, J., Ohan, V., Pollard, M.O., et al. (2021). Twelve years of SAMtools and BCFtools. GigaScience, 10, giab008.

Dawson, H.A., Quintella, B.R., Almeida, P.R., Treble, A.J. & Jolley, J.C. (2015). The ecology of larval and metamorphosing lampreys. In: Lampreys: Biology, conservation and control (ed. Docker, M.F.). Springer Netherlands, Dordrecht, pp. 75–137.

Docker, M.F., Bravener, G.A., Garroway, C.J., Hrodey, P.J., Hume, J.B., Johnson, N.S., et al. (2021). A review of sea lamprey dispersal and population structure in the Great Lakes and the implications for control. J. Gt. Lakes Res., 47, S549–S569.

Docker, M.F., Hume, J.B. & Clemens, B.J. (2015). Introduction: A surfeit of lampreys. In: Lampreys: Biology, conservation and control (ed. Docker, M.F.). Springer Netherlands, Dordrecht, pp. 1–34.

Docker, M.F. & Potter, I.C. (2019). Life history evolution in lampreys: Alternative migratory and feeding types. In: Lampreys: Biology, conservation and control (ed. Docker, M.F.). Springer Netherlands, Dordrecht, pp. 287–409.

Economidis, P.S., Kallianiotis, A. & Psaltopoulou, H. (1999). Two records of sea lamprey from the north Aegean Sea. J. Fish Biol., 55, 1114–1118.

Elliott, S., Deleys, N., Rivot, E., Acou, A., Réveillac, E. & Beaulaton, L. (2021). Shedding light on the river and sea lamprey in western European marine waters. Endanger. Species Res., 44, 409–419.

Foote, A.D., Vijay, N., Ávila-Arcos, M.C., Baird, R.W., Durban, J.W., Fumagalli, M., et al. (2016). Genome-culture coevolution promotes rapid divergence of killer whale ecotypes. Nat. Commun., 7, 11693.

Ford, M. (2024). *Petromyzon marinus*. IUCN Red List Threat. Species 2024 ET16781A58298056.

Frankham, R. (1995). Effective population size/adult population size ratios in wildlife: A review. Genet. Res., 66, 95–107.

Frankham R., Ballou, J.D., Briscoe, D.A. (2002). Introduction to conservation genetics Cambridge University Press, Cambridge.

Frichot, E. & François, O. (2015). LEA: An R package for landscape and ecological association studies. Methods Ecol. Evol., 6, 925–929.

Genner, M.J., Hillman, R., McHugh, M., Hawkins, S.J. & Lucas, M.C. (2012). Contrasting demographic histories of European and North American sea lamprey (*Petromyzon marinus*) populations inferred from mitochondrial DNA sequence variation. Mar. Freshw. Res., 63, 827.

Gibbard, P.L., Head, M.J., Walker, M.J.C., & the Subcommission on Quaternary Stratigraphy. (2010). Formal ratification of the Quaternary System/Period and the Pleistocene Series/Epoch with a base at 2.58 Ma. J. Quat. Sci., 25, 96–102.

Hardisty, M.W. & Potter, I.C. (1971). The general biology of adult lampreys. In: The biology of lampreys (eds. Hardisty M.W. & Potter I.C.). Academic Press, New York, pp. 127–206.

Hare, M.P., Nunney, L., Schwartz, M.K., Ruzzante, D.E., Burford, M., Waples, R.S., et al. (2011). Understanding and estimating effective population size for practical application in marine species management: Applying effective population size estimates to marine species management. Conserv. Biol., 25, 438–449.

Hartel, K.E., Halliwell, D.B., & Launer, A.E. (2002). Inland fishes of Massachusetts. Mass. Audubon Soc. Linc. Mass.

Hauser, L. & Carvalho, G.R. (2008). Paradigm shifts in marine fisheries genetics: Ugly hypotheses slain by beautiful facts. Fish Fish., 9, 333–362.

Hemmer-Hansen, J., Therkildsen, N.O. & Pujolar, J.M. (2014). Population genomics of marine fishes: Next-generation prospects and challenges. Biol. Bull., 227, 117–132.

Hess, J.E., Campbell, N.R., Close, D.A., Docker, M.F. & Narum, S.R. (2013). Population genomics of Pacific lamprey: Adaptive variation in a highly dispersive species. Mol. Ecol., 22, 2898–2916.

Hess, J.E., Delomas, T.A., Jackson, A.D., Kosinski, M.J., Moser, M.L., Porter, L.L., et al. (2022). Pacific lamprey translocations to the Snake River boost abundance of all life stages. Trans. Am. Fish. Soc., 151, 263–296.

Hewitt, G. (1999). Post-glacial re-colonization of European biota. Biol. J. Linn. Soc., 68, 87–112.

Hewitt, G. (2000). The genetic legacy of the Quaternary ice ages. Nature, 405, 907–913.

Hoban, S., Bruford, M.W., Funk, W.C., Galbusera, P., Griffith, M.P., Grueber, C.E., et al. (2021). Global commitments to conserving and monitoring genetic diversity are now necessary and feasible. BioScience, 71, 964–976.

Hogg, R., Coghlan, S.M. & Zydlewski, J. (2013). Anadromous sea lampreys recolonize a Maine coastal river tributary after dam removal. Trans. Am. Fish. Soc., 142, 1381–1394.

Holčík, J., Delić, A., Kučinić, M., Bukvić, V. & Vater, M. (2004). Distribution and morphology of the sea lamprey from the Balkan coast of the Adriatic Sea. J. Fish Biol., 64, 514–527.

Johnson, N.S., Buchinger, T.J. & Li, W. (2015). Reproductive ecology of lampreys. In: Lampreys: Biology, conservation and control (ed. Docker, M.F.). Springer Netherlands, Dordrecht, pp. 265–303.

Jombart, T. (2008). adegenet : A R package for the multivariate analysis of genetic markers. Bioinformatics, 24, 1403–1405.

Jombart, T. & Ahmed, I. (2011). adegenet 1.3-1 : New tools for the analysis of genome-wide SNP data. Bioinformatics, 27, 3070–3071.

Karachle, P.K. & Machias, A. (2014). First confirmed record of *Petromyzon marinus* Linnaeus, 1758 from West Greece. Cah. Biol. Mar., 55, 367–370.

King, J.J. & O' Gorman, N. (2018). Initial observations on feeding juvenile sea lamprey (*Petromyzon marinus* L.) in Irish lakes. Biol. Environ. Proc. R. Ir. Acad., 118B, 113.

Kynard, B. & Horgan, M. (2019). Long-term studies on restoration of Connecticut River anadromous sea lamprey, *Petromyzon marinus* Linnaeus 1758: Trend in annual adult runs, abundance cycle, and nesting. J. Appl. Ichthyol., 35, 1154–1163.

Lança, M.J., Machado, M., Mateus, C.S., Lourenço, M., Ferreira, A.F., Quintella, B.R., et al. (2014). Investigating population structure of sea lamprey (*Petromyzon marinus*, L.) in western Iberian Peninsula using morphological characters and heart fatty acid signature analyses. PLoS ONE, 9, e108110.

Lehner, F., Born, A., Raible, C.C. & Stocker, T.F. (2013). Amplified inception of European Little Ice Age by sea ice–ocean–atmosphere feedbacks. J. Clim., 26, 7586–7602.

Lehnert, S.J., Kess, T., Bentzen, P., Clément, M. & Bradbury, I.R. (2020). Divergent and linked selection shape patterns of genomic differentiation between European and North American Atlantic salmon (*Salmo salar*). Mol. Ecol., 29, 2160–2175.

Li, H. & Durbin, R. (2009). Fast and accurate short read alignment with Burrows–Wheeler transform. Bioinformatics, 25, 1754–1760.

Lin, B., Zhang, Z., Wang, Y., Currens, K.P., Spidle, A., Yamazaki, Y., et al. (2008). Amplified fragment length polymorphism assessment of genetic diversity in Pacific lampreys. North Am. J. Fish. Manag., 28, 1182–1193.

Livermore, J., Trainor, M., & Bednarski, M.S. (2017). Successful spawning of anadromous *Petromyzon marinus* L. (sea lamprey) in a restored stream channel following dam removal. Northeast. Nat., 24, 380–390.

Maitland, P.S., Renaud, C.B., Quintella, B.R., Close, D.A. & Docker, M.F. (2015). Conservation of native lampreys. In: Lampreys: Biology, conservation and control (ed. Docker, M.F.). Springer Netherlands, Dordrecht, pp. 375–428.

Mandić, A. & Petrić, L. (2021). Mediterranean Protected Areas in the Era of Overtourism. Springer, Cham.

Marsden, J.E. & Siefkes, M.J. (2019). Control of invasive sea lamprey in the Great Lakes, Lake Champlain, and Finger Lakes of New York. In: Lampreys: Biology, conservation and control (ed. Docker, M.F.). Springer Netherlands, Dordrecht, pp. 411–479.

Mateus, C., Rodríguez-Muñoz, R., Quintella, B., Alves, M. & Almeida, P. (2012). Lampreys of the Iberian Peninsula: Distribution, population status and conservation. Endanger. Species Res., 16, 183–198.

Mateus, C.S., Almeida, P.R., Mesquita, N., Quintella, B.R. & Alves, M.J. (2016). European lampreys: New insights on postglacial colonization, gene flow and speciation. PLoS ONE 11, e0148107.

Mateus, C.S., Docker, M.F., Evanno, G., Hess, J.E., Hume, J.B., Oliveira, I.C., et al. (2021). Population structure in anadromous lampreys: Patterns and processes. J. Gt. Lakes Res., 47, 38–58.

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., et al. (2010). The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res., 20, 1297–1303.

Mijangos, J.L., Gruber, B., Berry, O., Pacioni, C. & Georges, A. (2022). dartR v2: An accessible genetic analysis platform for conservation, ecology and agriculture. Methods Ecol. Evol., 13, 2150–2158.

Mizzan, L. & Vianello, C. (2007). Biodiversity of the lagoon of Venice and of the Venetian Northern Adriatic Coast. Boll. Mus. Civ. Storia Nat. Venezia, 58, 319–328.

Montanari, S.R., Hobbs, J.-P.A., Pratchett, M.S. & Van Herwerden, L. (2016). The importance of ecological and behavioural data in studies of hybridisation among marine fishes. Rev. Fish Biol. Fish., 26, 181–198.

Morin, P.A., Luikart, G., Wayne, R.K., & The Snp Workshop Group. (2004). SNPs in ecology, evolution and conservation. Trends Ecol. Evol., 19, 208–216.

Morman, R.H. (1979). Distribution and ecology of lampreys in the Lower Peninsula of Michigan, 1957–75 (Technical Report No. 33). Great Lakes Fishery Commission, Ann Arbor.

Moser, M.L., Almeida, P.R., King, J.J. & Pereira, E. (2021). Passage and freshwater habitat requirements of anadromous lampreys: Considerations for conservation and control. J. Gt. Lakes Res., 47, S147–S158.

Murauskas, J.G., Orlov, A.M., Keller, L., Maznikova, O.A. & Glebov, I.I. (2019). Transoceanic migration of Pacific lamprey, *Entosphenus tridentatus*. J. Ichthyol., 59, 280–282.

NatureServe (2024). NatureServe network biodiversity location data accessed through NatureServe explorer [web application]. http://explorer.natureserve.org

Neely, K.L., Lewis, C.L., O'Neil, K., Woodley, C.M., Moore, J., Ransom, Z., et al. (2021). Saving the last unicorns: The genetic rescue of Florida's pillar corals. Front. Mar. Sci., 8, 657429.

Nikolic, N., Liu, S., Jacobsen, M.W., Jónsson, B., Bernatchez, L., Gagnaire, P., et al. (2020). Speciation history of European (*Anguilla anguilla*) and American eel (*A. rostrata*), analysed using genomic data. Mol. Ecol., 29, 565–577.

Nilsson, J., Gross, R., Asplund, T., Dove, O., Jansson, H., Kelloniemi, J., et al. (2001). Matrilinear phylogeography of Atlantic salmon (*Salmo salar* L.) in Europe and postglacial colonization of the Baltic Sea area. Mol. Ecol., 10, 89–102.

Novak, B.J., Phelan, R. & Weber, M. (2021). U.S. conservation translocations: Over a century of intended consequences. Conserv. Sci. Pract., 3, e394.

Palsboll, P., Berube, M. & Allendorf, F. (2007). Identification of management units using population genetic data. Trends Ecol. Evol., 22, 11–16.

Palstra, F.P. & Ruzzante, D.E. (2008). Genetic estimates of contemporary effective population size: What can they tell us about the importance of genetic stochasticity for wild population persistence? Mol. Ecol., 17, 3428–3447.

Patarnello, T., Volckaert, F.A.M.J. & Castilho, R. (2007). Pillars of Hercules: Is the Atlantic–Mediterranean transition a phylogeographical break? Mol. Ecol., 16, 4426–4444.

Pereira, A.M., Jonsson, B., Johannsson, M., Robalo, J.I. & Almada, V.C. (2012). Icelandic lampreys (*Petromyzon marinus*): where do they come from? Ichthyol. Res., 59, 83–85.

Petetta, A., Righi, T., Splendiani, A., Virgili, M., Giovannotti, M., Lucchetti, A., et al. (2019). A new mtDNA control region haplotype from sea lamprey (*Petromyzon marinu*s Linnaeus, 1758: Petromyzontiformes) collected off the Italian Adriatic coast. Mar. Biodivers., 49, 2917–2920.

Piavis, G.W., Howell, J.H. & Smith, A.J. (1970). Experimental hybridization among five species of lampreys from the Great Lakes. Copeia, 1970, 29.

Potter, I.C., Gill, H.S., Renaud, C.B. & Haoucher, D. (2015). The taxonomy, phylogeny, and distribution of lampreys. In: Lampreys: Biology, conservation and control (ed. Docker, M.F.). Springer Netherlands, Dordrecht, pp. 35–73.

Potter, I.C. & Hilliard, R.W. (1987). A proposal for the functional and phylogenetic significance of differences in the dentition of lampreys (Agnatha: Petromyzontiformes). J. Zool., 212, 713–737.

R Core Team (2023). R: A language and environment for statistical computing.

Rafrafi-Nouira, S., El Kamel-Moutalibi, O., Boumaiza, M., Reynaud, C., & Capapé, C. (2015). On the occurrence of sea lamprey *Petromyzon marinus* (Agnatha: Petromyzonidae) in Tunisian waters (Central Mediterranean Sea). Cah. Biol. Mar., 56, 73–76.

Renaud, C.B. & Cochran, P.A. (2019). Post-metamorphic feeding in lampreys. In: Lampreys: Biology, conservation and control (ed. Docker, M.F.). Springer Netherlands, Dordrecht, pp. 247–285.

Rodríguez-Muñoz, R., Waldman, J.R., Grunwald, C., Roy, N.K. & Wirgin, I. (2004). Absence of shared mitochondrial DNA haplotypes between sea lamprey from North American and Spanish rivers. J. Fish Biol., 64, 783–787.

Rose, G.A., Marteinsdóttir, G. & Godø, O. (2019). Exploitation: Cod is fish and fish is cod. In: Atlantic Cod (ed. Rose, G.A.). Wiley, pp. 287–336.

Rougemont, Q. & Bernatchez, L. (2018). The demographic history of Atlantic salmon (*Salmo salar*) across its distribution range reconstructed from approximate Bayesian computations: Atlantic salmon history and linked selection. Evolution, 72, 1261–1277.

Rougemont, Q., Dolo, V., Oger, A., Besnard, A.-L., Huteau, D., Coutellec, M.-A., et al. (2021). Riverscape genetics in brook lamprey: Genetic diversity is less influenced by river fragmentation than by gene flow with the anadromous ecotype. Heredity, 126, 235–250.

Sabatino, S.J., Faria, R. & Alexandrino, P.B. (2022). Genetic structure, diversity, and connectivity in anadromous and freshwater *Alosa alosa* and *A. fallax.* Mar. Biol., 169, 2.

Santiago, E., Novo, I., Pardiñas, A.F., Saura, M., Wang, J. & Caballero, A. (2020). Recent demographic history inferred by high-resolution analysis of linkage disequilibrium. Mol. Biol. Evol., 37, 3642–3653.

Sard, N.M., Smith, S.R., Homola, J.J., Kanefsky, J., Bravener, G., Adams, J.V., et al. (2020). RAPTURE (RAD capture) panel facilitates analyses characterizing sea lamprey reproductive ecology and movement dynamics. Ecol. Evol., 10, 1469–1488.

Scribner, K.T., Page, K.S. & Bartron, M.L. (2000). Hybridization in freshwater fishes: A review of case studies and cytonuclear methods of biological inference. Rev. Fish Biol. Fish., 10, 293–323.

Smith, J.J. & Keinath, M.C. (2015). The sea lamprey meiotic map improves resolution of ancient vertebrate genome duplications. Genome Res., 25, 1081–1090.

Spice, E.K., Goodman, D.H., Reid, S.B. & Docker, M.F. (2012). Neither philopatric nor panmictic: microsatellite and mtDNA evidence suggests lack of natal homing but limits to dispersal in Pacific lamprey. Mol. Ecol., 21, 2916–2930.

Terhorst, J., Kamm, J.A. & Song, Y.S. (2017). Robust and scalable inference of population history from hundreds of unphased whole genomes. Nat. Genet., 49, 303–309.

Thessalou-Legaki, M., Aydogan, O., Bekas, P., Bilge, G., Boyaci, Y.O., Brunelli, E., et al. (2012). New Mediterranean biodiversity records (December 2012). Mediterr. Mar. Sci., 13, 312.

Timoshevskaya, N., Eşkut, K.İ., Timoshevskiy, V.A., Robb, S.M.C., Holt, C., Hess, J.E., et al. (2023). An improved germline genome assembly for the sea lamprey *Petromyzon marinus* illuminates the evolution of germline-specific chromosomes. Cell Rep., 42, 112263.

Van Leeuwen, T.E., Cote, D., Pretty, C., Townley, J., Poole, R., Dempson, B., et al. (2021). Incursions of sea lamprey, *Petromyzon marinus*, and striped bass, *Morone saxatilis*, in Labrador waters: Episodic events or evidence of a northward range expansion? J. Fish Biol., 99, 2066–2070.

Waldman, J., Grunwald, C. & Wirgin, I. (2008). Sea lamprey P*etromyzon marinus*: An exception to the rule of homing in anadromous fishes. Biol. Lett., 4, 659–662.

Waldman, J.R., Grunwald, C. & Wirgin, I. (2006). Evaluation of the native status of sea lampreys in Lake Champlain based on mitochondrial DNA sequencing analysis. Trans. Am. Fish. Soc., 135, 1076–1085.

Wang, J. & Whitlock, M.C. (2003). Estimating effective population size and migration rates from genetic samples over space and time. Genetics, 163, 429–446.

Waples, R.K., Larson, W.A. & Waples, R.S. (2016). Estimating contemporary effective population size in non-model species using linkage disequilibrium across thousands of loci. Heredity, 117, 233–240.

Weir, B.S. & Cockerham, C.C. (1984). Estimating F-statistics for the analysis of population structure. Evolution, 38, 1358.

WGS500 Consortium, Rimmer, A., Phan, H., Mathieson, I., Iqbal, Z., Twigg, S.R.F., et al. (2014). Integrating mapping-, assembly- and haplotype-based approaches for calling variants in clinical sequencing applications. Nat. Genet., 46, 912–918.

Wilkie, M.P., Johnson, N.S. & Docker, M.F. (2022). Invasive species control and management: The sea lamprey story. In: Fish Physiology. Elsevier, pp. 489–579.

Xuereb, A., D'Aloia, C.C., Daigle, R.M., Andrello, M., Dalongeville, A., Manel, S., et al. (2019). Marine conservation and marine protected areas. In: Population genomics: Marine organisms, population genomics (eds. Oleksiak, M.F. & Rajora, O.P.). Springer International Publishing, Cham, pp. 423–446.

Yamasaki, Y.Y., Kakioka, R., Takahashi, H., Toyoda, A., Nagano, A.J., Machida, Y., et al. (2020). Genome-wide patterns of divergence and introgression after secondary contact between *Pungitius* sticklebacks. Philos. Trans. R. Soc. B Biol. Sci., 375, 20190548.

Zimmerman, S.J., Aldridge, C.L. & Oyler-McCance, S.J. (2020) An empirical comparison of population genetic analyses using microsatellite and SNP data for a species of conservation concern. BMC Genomics, 21, 382.

Table 3.1. Anadromous sea lamprey samples and associated metadata.

Sampling Site	Location	Abbreviation	Coordinates	Samples	Collection Date	Life Stage	Preservation Method	Collaborator	Affiliation
Close to Vestra-Horn	Iceland	ICE_67 ICE_68	64.162500, -14.476333	2	2009	Adult	Tissue - ethanol	Benóný Jónsson, Magnús Jóhannsson	Hafrannsóknastofnun (MFRI)
Close to Vestman Islands	Iceland	ICE_69	63.347633, -20.087850	1	2009	Adult	Tissue - ethanol	Benóný Jónsson, Magnús Jóhannsson	Hafrannsóknastofnun (MFRI)
Ytri Ranga River	Iceland	ICE_66	63.818367, -20.418817	1	2009	Adult	Tissue - ethanol	Benóný Jónsson, Magnús Jóhannsson	Hafrannsóknastofnun (MFRI)

At sea Heradsfloi	Iceland	ICE_70 - ICE_74	65.657933, -14.135033	5	2012	Adult	Tissue - ethanol	Benóný Jónsson, Magnús Jóhannsson	Hafrannsóknastofnun (MFRI)
River Rhine	Germany	PMGE2020	48.83316932, 8.10988970295832	8	2020	Adult & larvae	Fin clips - ethanol	Jan Baer	Landwirtschaftliches Zentrum Baden-Württemberg (LAZBW)
Lough Conn – Moy*	Ireland	Ireland_01 - Ireland_09	N/A	9	2010, 2011, 2016, 2017	Adult	Fin clips - ethanol	Fiona Bracken	University College Dublin
River Mulkear - Shannon	Ireland	Ireland_11 - Ireland_27	N/A	10	2010, 2011	Adult	Fin clips - ethanol	Fiona Bracken	University College Dublin
Lough Derg - Shannon*	Ireland	Ireland_28 - Ireland_37	N/A	8	2010, 2011, 2017	Adult	Fin clips - ethanol	Fiona Bracken	University College Dublin
Oir River	France Atlantic	SEL	N/A	7	2015, 2017, 2018	Adult	Fin clips - ethanol	Guillaume Evanno	Institut national de recherche pour l'agriculture, l'alimentation et l'environnement (INRAE)
Scorff River	France Atlantic	SC	N/A	8	2018	Adult	Fin clips - ethanol	Guillaume Evanno	Institut national de recherche pour l'agriculture, l'alimentation et l'environnement (INRAE)
Loire River	France Atlantic	PEMA	N/A	6	2013	Adult	Fin clips - ethanol	Guillaume Evanno	Institut national de recherche pour l'agriculture, l'alimentation et l'environnement (INRAE)
Vouga River Basin	Portugal	VOU	N/A	10	2021	Adult	Fin clips - ethanol	Catarina Mateus	Universidade de Évora
Mondego River Basin	Portugal	MON	N/A	8	2020	Adult	Fin clips - ethanol	Catarina Mateus	Universidade de Évora
Tejo River Basin	Portugal	TEJO	N/A	9	2021	Adult	Fin clips - ethanol	Catarina Mateus	Universidade de Évora
Guadiana River Basin	Portugal	GUA	N/A	10	2020, 2021	Adult	Fin clips - ethanol	Catarina Mateus	Universidade de Évora
Hogvadsan	Sweden	SWE	N/A	7	2021	Adult	Fin clips - ethanol	Mikael Svensson, Elisabeth Thysell	Sveriges lantbruksuniversitet (SLU), Länsstyrelsen i Hallands län
Rhône Estuary	France Mediterranean	FRMED	N/A	6	2003, 2010, 2015	Adult & NA	Fin clips - ethanol	Guillaume Evanno	Institut national de recherche pour l'agriculture, l'alimentation et l'environnement (INRAE)
Connecticut River	Massachusetts	CTR	41.28477, -72.342589	12	2019	Adult	Fin clips – frozen carcasses	Ted Castro-Santos	U.S. Geological Survey (USGS)

Richibucto River	New Brunswick	RI	46.692742, -64.842909	12	2019	Larvae	Fin clips – frozen carcasses	Mike Wilkie	Wilfrid Laurier University
Deer Creek	Maryland	DR2021	39.6748228, -76.4511499	2	2021	Adult	Fin clips - ethanol	Thomas Evans	St. Mary's College of Maryland
Falling Branch	Maryland	FB2021	39.6909291, -76.4265029	2	2021	Adult	Fin clips - ethanol	Thomas Evans	St. Mary's College of Maryland
Little Patuxent River	Maryland	PR2021	39.134949, -76.82557	4	2021	Larvae	Fin clips - ethanol	Thomas Evans	St. Mary's College of Maryland
Pee Dee River**	North Carolina	NCSM_69176 NCSM_69181	34.922800, -80.018530 34.856270, -79.916770	2	2011	NA	Tissue - ethanol	Gabriela M.Hogue, Michael Fisk	North Carolina Museum of Natural Sciences, NC Wildlife Resources Commission
Neuse River**	North Carolina	NCSM_46040	35.349, -78.1501	1	2007	NA	Tissue - ethanol	Gabriela M. Hogue, Michael Fisk	North Carolina Museum of Natural Sciences, NC Wildlife Resources Commission
Roanoke River**	North Carolina	NCSM_38060 NCSM_105133 NCSM_97531 NCSM_97532	36.328700, -77.578100 36.477120, -77.639720 36.481510 to 36.468700, -77.643820 to -77.636830	4	2018, 2019, 2024	Adult & NA	Tissue - ethanol	Gabriela M. Hogue, Michael Fisk	North Carolina Museum of Natural Sciences, NC Wildlife Resources Commission
Pamlico River, Tar River**	North Carolina	NCSM 53260	35.955400, -77.780700	1	2009	Adult	Tissue - ethanol	Gabriela M.Hogue, Michael Fisk	North Carolina Museum of Natural Sciences, NC Wildlife Resources Commission

* Locations at which feeding in fresh water prior to outmigration has been observed, ** Samples from museum collection (North Carolina Museum of Natural Sciences Ichthyology Collection).

Table 3.2. Mean pairwise F_{ST} estimates as per Weir & Cockerham (1984) (window size 20kb) between distinct sampling locations. Significantly different values in bold.

Mean FST	Ireland	France	Germany	Portugal	Iceland	Sweden	Maryland	Massachusetts	N. Carolina	N. Brunswick
Ireland	-	0.0027604	0.0017698	0.00036153	-0.00051566	-0.0013132	**0.24939**	**0.23786**	**0.25022**	**0.23839**
France	0.0027604	-	0.0024385	0.0019974	0.00086998	0.00052472	**0.24012**	**0.22899**	**0.24067**	**0.22942**
Germany	0.0017698	0.0024385	-	0.0015472	-0.00049131	-0.0018219	**0.19562**	**0.19171**	**0.19536**	**0.19193**

Portugal	0.00036153	0.0019974	0.0015472	-	-0.00065924	-0.0015786	**0.26603**	**0.25382**	**0.26697**	**0.25429**
Iceland	-0.00051566	0.00086998	-0.00049131	-0.00065924	-	-0.0022672	**0.19976**	**0.19465**	**0.19965**	**0.1952**
Sweden	-0.0013132	0.00052472	-0.0018219	-0.0015786	-0.0022672	-	**0.1902**	**0.18723**	**0.18985**	**0.18759**
Maryland	**0.24939**	**0.24012**	**0.19562**	**0.26603**	**0.19976**	**0.1902**	-	0.0023248	-0.00017578	0.0021566
Massachusetts	**0.23786**	**0.22899**	**0.19171**	**0.25382**	**0.19465**	**0.18723**	0.0023248	-	-0.00076697	-0.0010563
N. Carolina	**0.25022**	**0.24067**	**0.19536**	**0.26697**	**0.19965**	**0.18985**	-0.00017578	-0.00076697	-	-0.0014812
N. Brunswick	**0.23839**	**0.22942**	**0.19193**	**0.25429**	**0.1952**	**0.18759**	0.0021566	-0.0010563	-0.0014812	-

Table 3.3. Mean pairwise F_{ST} estimates as per Weir & Cockerham (1984) (window size 20kb) between the three different sub-sampling locations in Ireland.

Ireland	Lough Derg (Shannon)	Lough Conn (Moy)	River Mulkear (Shannon)
Lough Derg (Shannon)	-	0.001	-0.002
Lough Conn (Moy)	0.001	-	0.002
River Mulkear (Shannon)	-0.002	0.002	-

Table 3.4. Mean pairwise F_{ST} estimates as per Weir & Cockerham (1984) (window size 20kb) between the three different sub-sampling locations in France.

France	Loire River	Oir River	Scorff River
Loire River	-	-0.004	-0.003
Oir River	-0.004	-	-0.003
Scorff River	-0.003	-0.003	-

Table 3.5. Mean pairwise F_{ST} estimates as per Weir & Cockerham (1984) (window size 20kb) between the four different sub-sampling locations in Portugal.

Portugal	Guadiana River basin	Mondego River basin	Tejo River basin	Vouga River basin
Guadiana River basin	-	-0.002	-0.002	-0.001
Mondego River basin	-0.002	-	-0.002	-0.001

Tejo River basin	-0.002	-0.002	-	-0.002
Vouga River basin	-0.001	-0.001	-0.002	-

Table 3.6. Nucleotide diversity estimates (window size 20kb) across all distinct sampling and sub-sampling locations.

Location	Mean PI
West Atlantic	1.125e-04
East Atlantic	7.643e-05
Europe excluding Mediterranean Indiv.	7.632e-05
Mediterranean (French coast)	7.898e-05
Ireland (all locations)	7.702e-05
Lough Derg Shannon (Ireland)	7.902e-05
Lough Conn Moy (Ireland)	7.792e-05
River Mulkear Shannon (Ireland)	7.822e-05
France (all Atlantic locations)	7.821e-05
Loire River (France)	8.162e-05
Oir River (France)	8.056e-05
Scorff River (France)	7.997e-05

Location	Mean PI
Germany	7.935e-05
Portugal (all locations)	7.678e-05
Guadiana River basin (Portugal)	7.845e-05
Mondego River basin (Portugal)	7.874e-05
Tejo River basin (Portugal)	7.866e-05
Vouga River basin (Portugal)	7.843e-05
Iceland	7.871e-05
Maryland	1.109e-04
Massachusetts	1.137e-04
North Carolina	1.120e-04
New Brunswick	1.136e-04
Sweden	7.957e-05

Table 3.7. Contemporary effective population size (Ne) estimates as per Waples *et al.* (2016) across all distinct sampling locations. The mean across all five runs per location is reported in the main text.

Location	Ne Run 1	Ne Run 2	Ne Run 3	Ne Run 4	Ne Run 5
Mediterranean (French coast)	24,18239	24,53278	24,68503	24,81832	24,0143
Ireland	1258,501	819,7898	1679,848	860,6329	1309,688
France (Atlantic)	2901,156	1982,278	2102,389	2225,005	2088,504

Germany	3,060746	2,834029	2,962041	2,759615	2,834494
Portugal	1855,5	4765,583	6594,048	16524,39	5247,671
Iceland	328,2581	380,2433	313,0023	451,6325	319,242
Maryland	38,82678	35,60967	38,23725	38,4545	38,50506
Massachusetts	8053,25	14369,26	21434,43	2178,514	2754,337
North Carolina	152,9915	141,1976	149,8796	152,4433	156,5964
New Brunswick	Inf	Inf	Inf	Inf	Inf
Sweden	63,72907	63,68867	60,05021	54,59841	64,23507

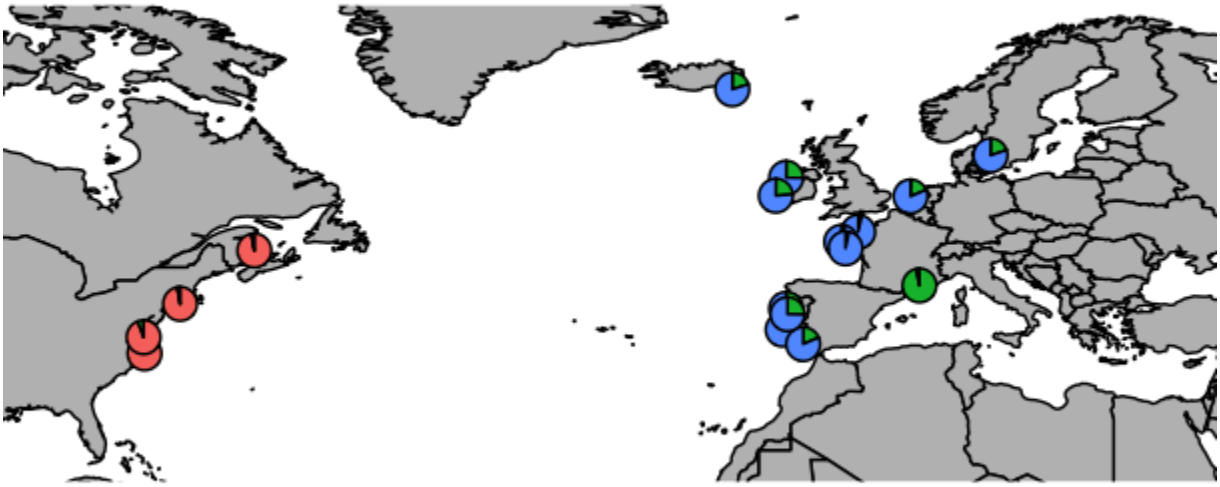

Figure 3.1. Map of ancestry coefficients estimated using snmf in LEA (Frichot & François 2015). Analysis was run for all sea lamprey individuals (n = 155) for an assumed number of genetically distinct populations K = 3 here represented by different colors. Each pie represents a geographically distinct sampling location. When metadata associated with samples did not include coordinates, or coordinates were in freshwater, coordinates were taken from the drainage basin of the water body individuals were

sampled from. In the case of Iceland, where sea lamprey are not known to spawn and individuals were caught opportunistically (mostly at sea), coordinates were averaged across individuals.

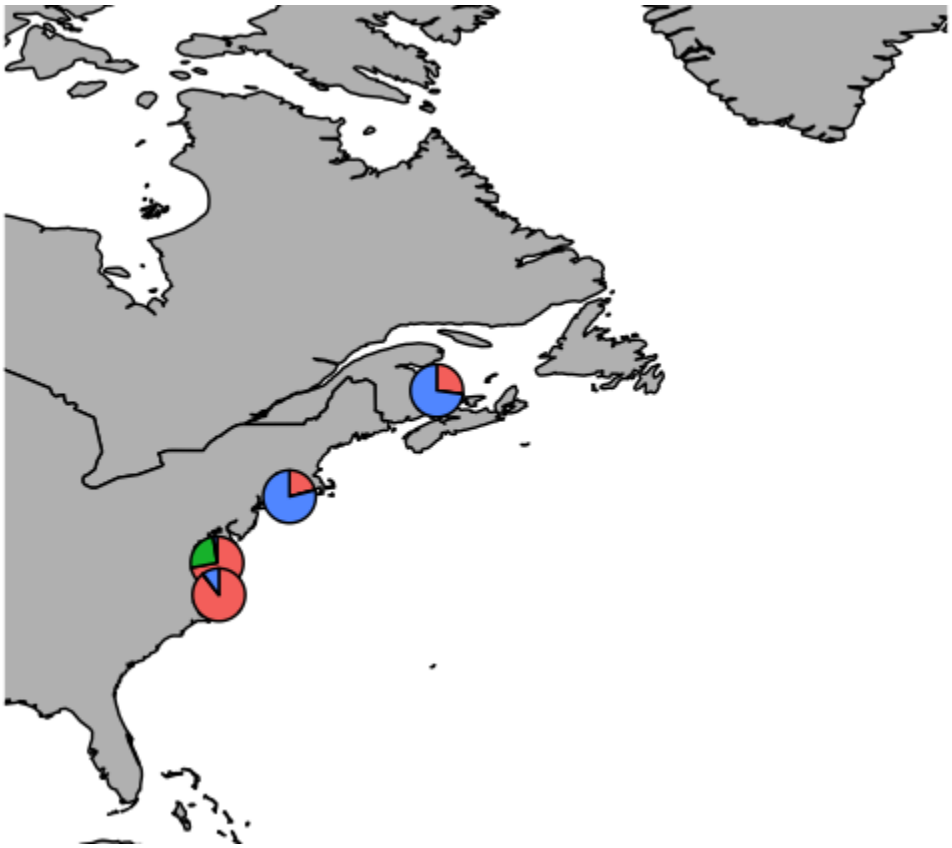

Figure 3.2. Map of ancestry coefficients estimated using snmf in LEA (Frichot & François 2015). Analysis was run for the West Atlantic North American coast sea lamprey individuals (n = 40) for an assumed number of genetically distinct populations K = 3 here represented by different colors. Each pie represents a geographically distinct sampling location. When metadata associated with samples did not include coordinates, or coordinates were in freshwater, coordinates were taken from the drainage basin of the water body individuals were sampled from.

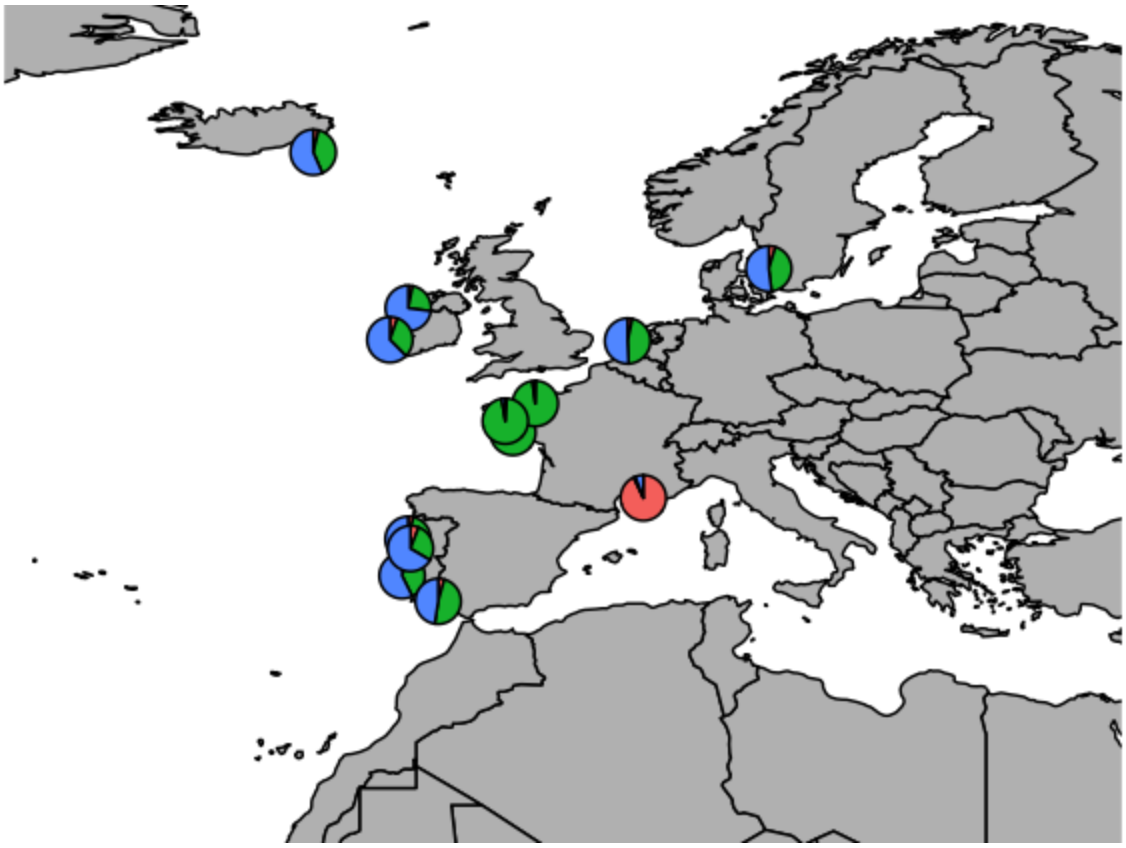

Figure 3.3. Map of ancestry coefficients estimated using snmf in LEA (Frichot & François 2015). Analysis was run for East Atlantic and Mediterranean European sea lamprey individuals (n = 115) for an assumed number of genetically distinct populations K = 3 here represented by different colors. When metadata associated with samples did not include coordinates, or coordinates were in freshwater, coordinates were taken from the drainage basin of the water body individuals were sampled from. In the case of Iceland, where sea lamprey are not known to spawn and individuals were caught opportunistically (mostly at sea), coordinates were averaged across individuals.

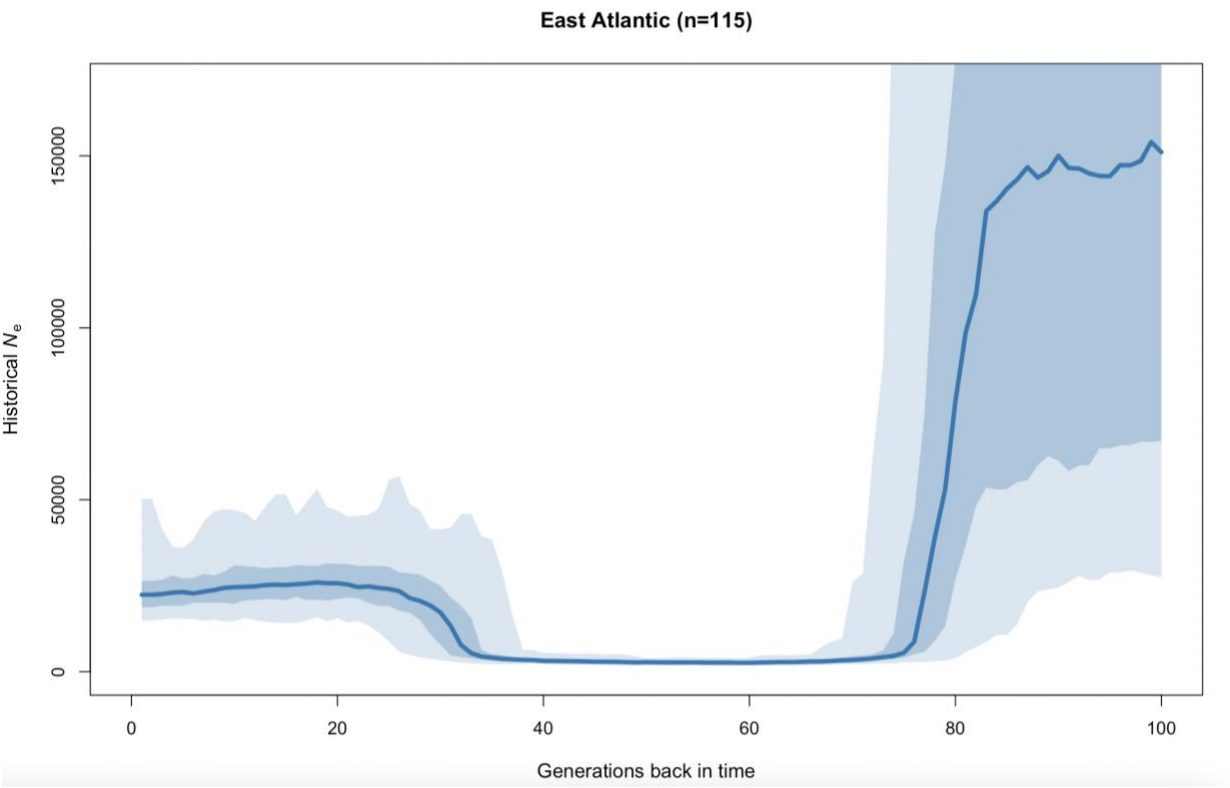

Figure 3.4. Plot of 100 replicate model runs reconstructing recent demographic history for the East Atlantic (115 individuals including the French Mediterranean coast), showing the trajectory of effective population size for the last 100 generations (900 years). Analysis undertaken in GONE (Santiago *et al.* 2020). Effective population size on the y axis and generations back in time on the x axis. Dark blue shading represents the 90% confidence interval and light blue shading represents the 95% confidence interval.

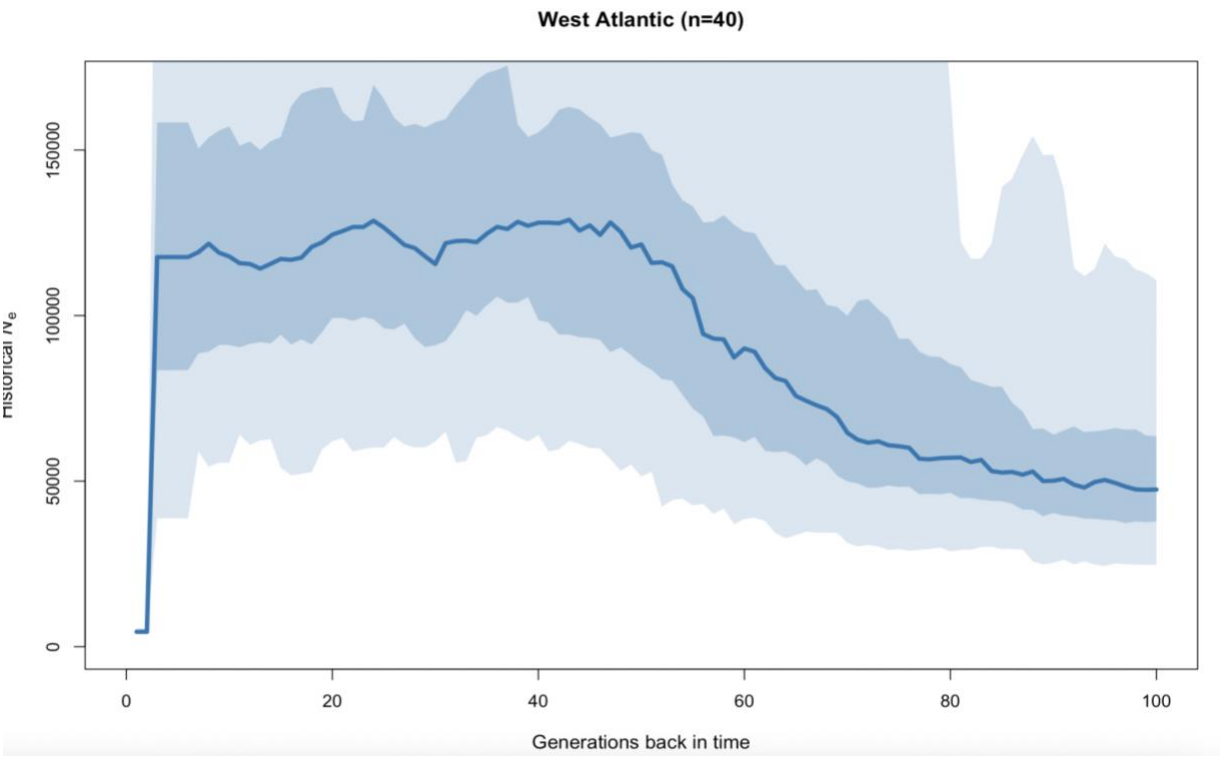

Figure 3.5. Plot of 100 replicate model runs reconstructing recent demographic history for the West Atlantic (40 individuals), showing the trajectory of effective population size for the last 100 generations (900 years). Analysis undertaken in GONE (Santiago *et al.* 2020). Effective population size on the y axis and generations back in time on the x axis. Dark blue shading represents the 90% confidence interval and light blue shading represents the 95% confidence interval.

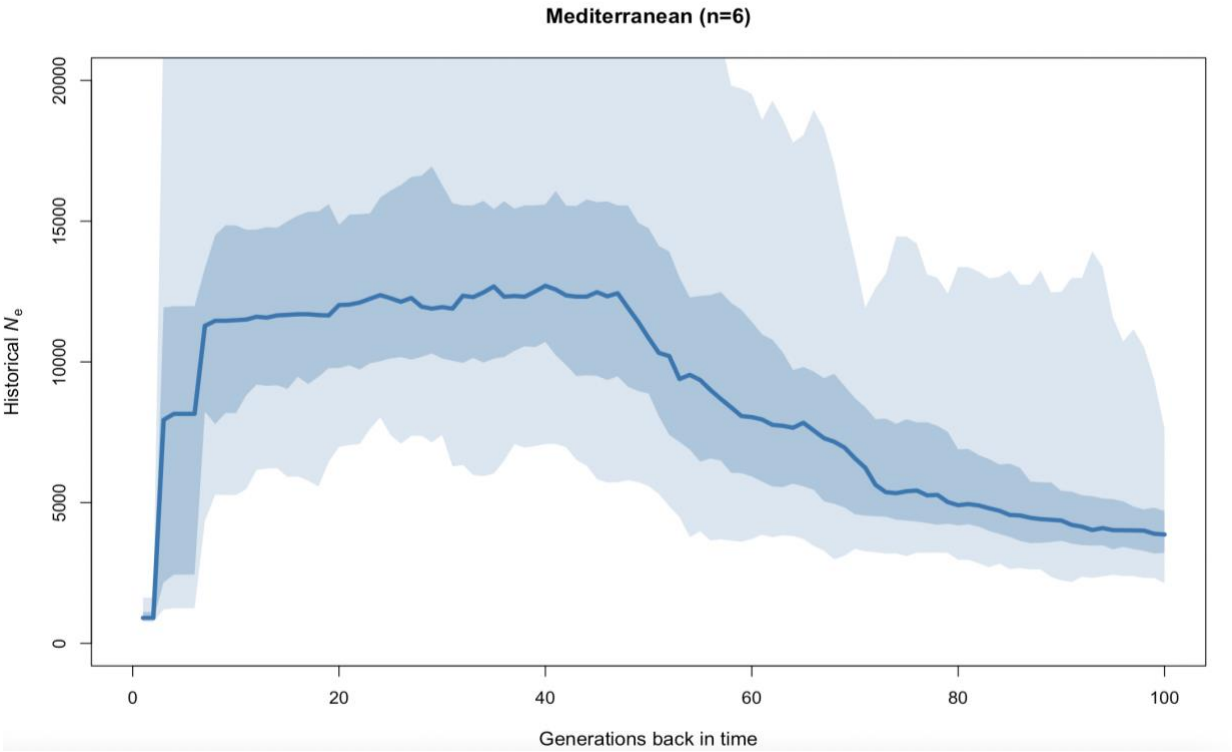

Figure 3.6. Plot of 100 replicate model runs reconstructing recent demographic history for the Mediterranean Sea (French coast - 6 individuals) separately, showing the trajectory of effective population size for the last 100 generations (900 years). Analysis undertaken in GONE (Santiago *et al.* 2020). Effective population size on the y axis and generations back in time on the x axis. Dark blue shading represents the 90% confidence interval and light blue shading represents the 95% confidence interval.

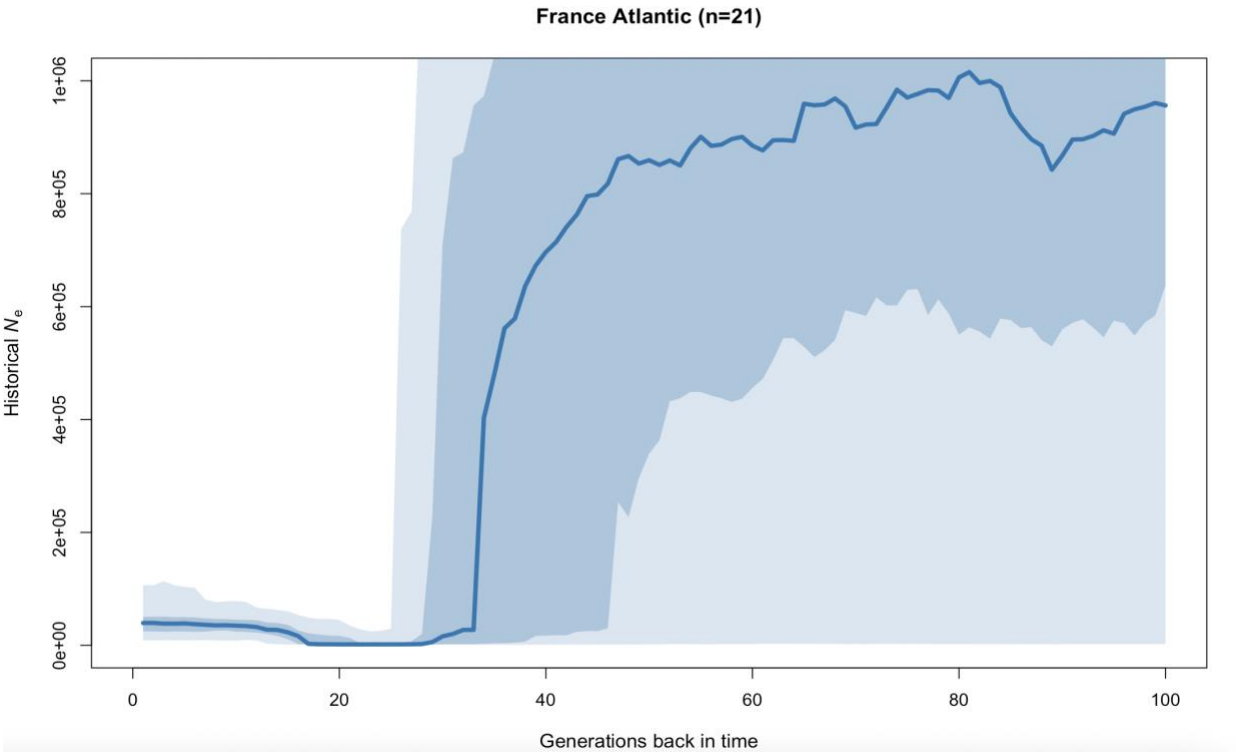

Figure 3.7. Plot of 100 replicate model runs reconstructing recent demographic history for the French Atlantic coast (21 individuals) separately, showing the trajectory of effective population size for the last 100 generations (900 years). Analysis undertaken in GONE (Santiago *et al.* 2020). Effective population size on the y axis and generations back in time on the x axis. Dark blue shading represents the 90% confidence interval and light blue shading represents the 95% confidence interval.

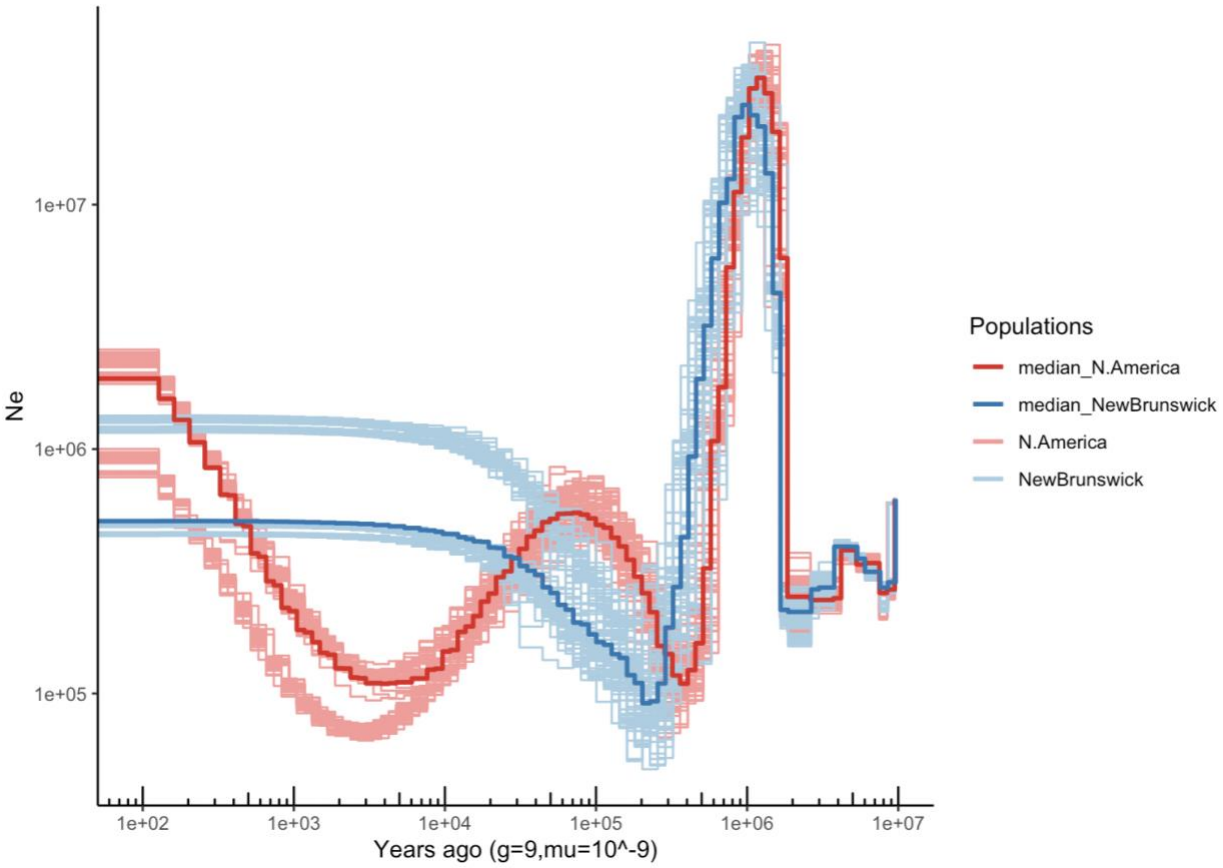

Figure 3.8. Plot of 100 replicate model runs reconstructing deep demographic history for the entire West Atlantic North American coast (n = 40) and individuals from New Brunswick used to estimate the transatlantic split (n = 10). The bold lines represent the median across model runs. Analysis undertaken in SMC++ (Terhorst *et al.* 2017). Effective population size on the y axis. Years back in time based on a per generation mutation rate of 10^{-9} on the x axis. Models estimated across a timescale of 10 - 10^{6} years ago.

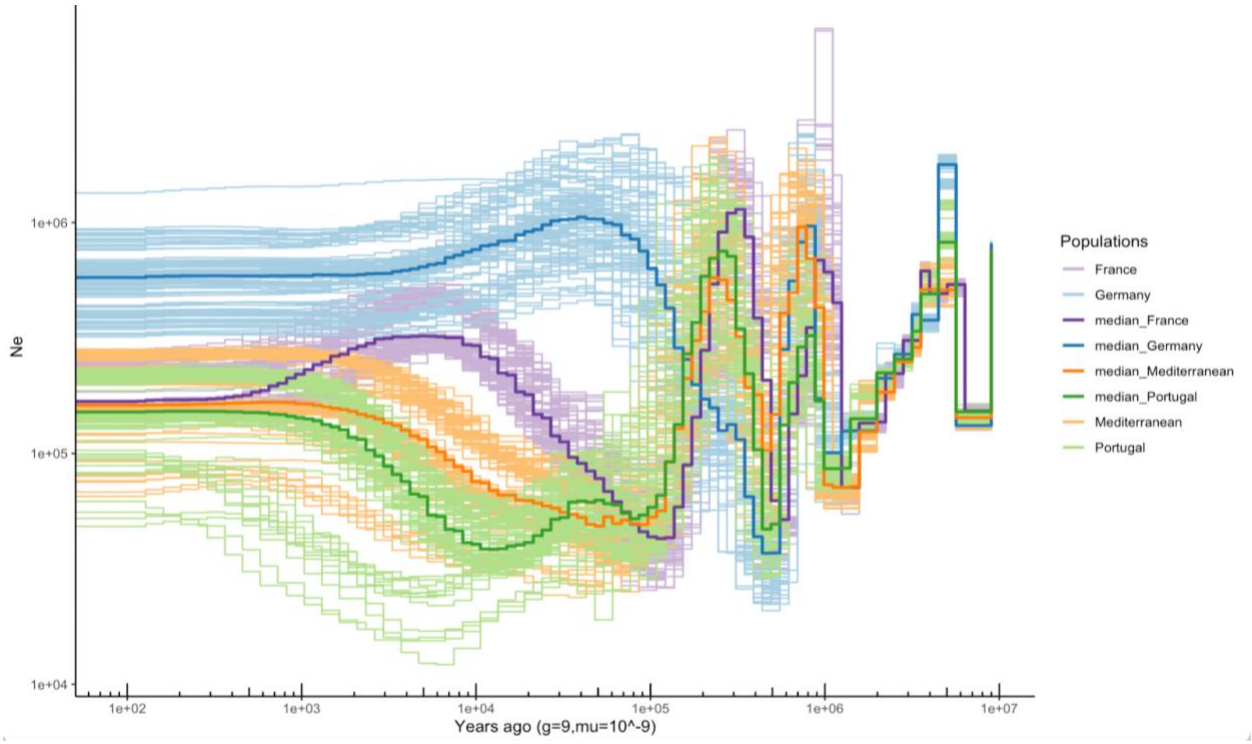

Figure 3.9. Plot of 100 replicate model runs reconstructing deep demographic history for individuals from Germany (n = 8), Portugal (10 individuals used to estimate the transatlantic split), the French Atlantic (n = 21) and Mediterranean (n = 6) coasts as representatives of the East Atlantic European and Mediterranean region. The bold lines represent the median across model runs. Analysis undertaken in SMC++ (Terhorst *et al.* 2017). Effective population size on the y axis. Years back in time based on a per generation mutation rate of 10^{-9} on the x axis. Models estimated across a timescale of 10 - 10^{6} years ago.

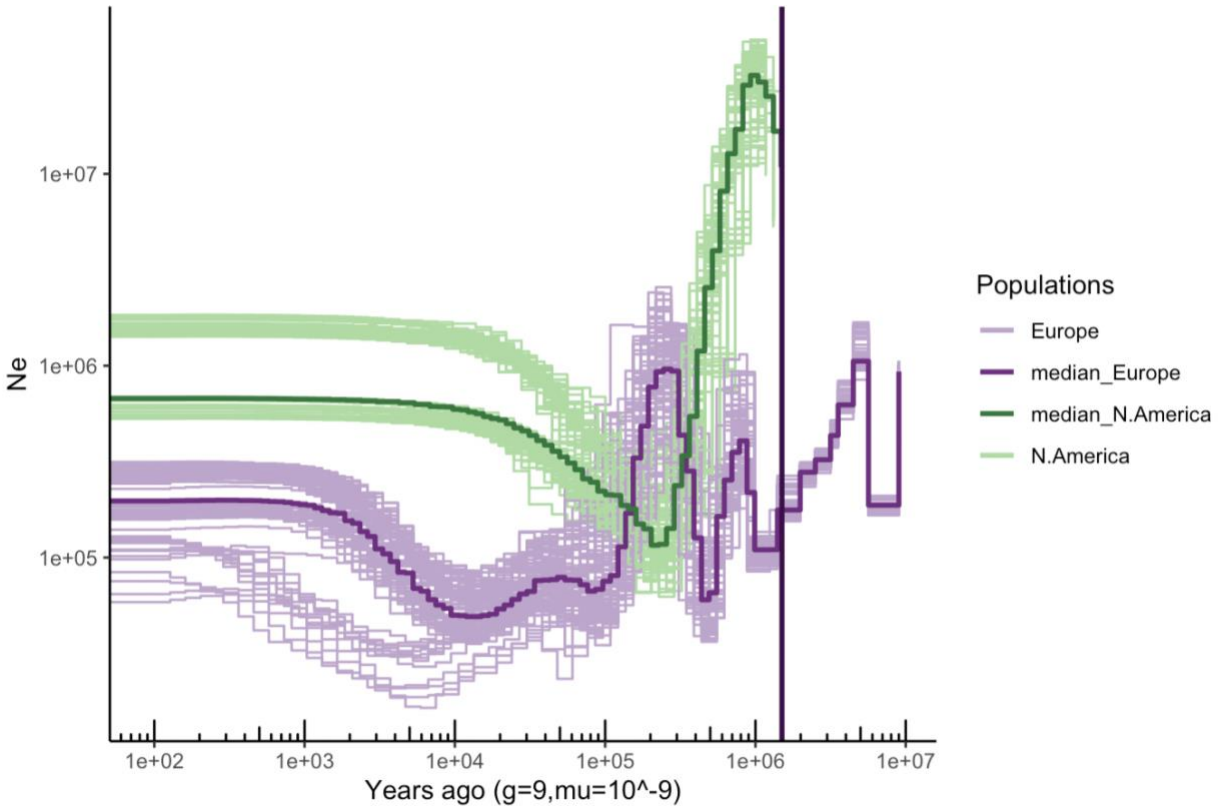

Figure 3.10. Plot of 100 replicate model runs reconstructing the transatlantic split using 10 individuals from New Brunswick and 10 individuals from Portugal representing the West and East Atlantic coasts respectively. The bold lines represent the median across model runs. Analysis undertaken in SMC++ (Terhorst *et al.* 2017). Effective population size on the y axis. Years back in time based on a per generation mutation rate of 10^{-9} on the x axis. Models estimated across a timescale of 10 - 10^{6} years ago. The timing of the transatlantic split is represented by the black vertical line.

CHAPTER 4: MALADAPTATION RISK ACROSS THE ANADROMOUS SEA LAMPREY NATIVE RANGE

Karachaliou E. [1], Suurväli J.[1], de Greef E. [1], Garroway C.J. [1], Docker M.F. [1]

[1] Department of Biological Sciences, University of Manitoba, 50 Sifton Rd, Winnipeg, MB, R3T 2N2, Canada

Abstract

Sea lamprey (*Petromyzon marinus*), subject to habitat degradation, barriers to migration, and the increasing effects of climate change, is of conservation concern across its native range. I use genome-wide information to determine adaptation to local environmental conditions and assess future vulnerability under climate change. Whole genome resequencing was undertaken for 155 individuals from across the sea lamprey native range. I initially conducted a scan for local adaptation to differences in temperature, oxygen, pH, salinity and urbanization. I then calculated genetic offsets to predict per-location selection pressure intensity under future climate change. I found the main driver of local adaptation in sea lamprey to be differences in pH levels. Using genetic offsets, I forecasted locations at which sea lamprey are most likely to experience intense selection pressure and are therefore more vulnerable to future climate change. I discuss conservation priorities based on both contemporary sea lamprey status and predicted future vulnerability.

Introduction

Strong effects of climate change are expected in North Atlantic, Mediterranean and Baltic Sea ecosystems, predominantly related to increasing temperature (Andersson *et al.* 2015; Lejeusne *et al*. 2010; Poloczanska *et al.* 2013). Regionally the effects of climate change can vary. For example, in the Mediterranean predicted increases in temperature and extreme weather events are already underway (reviewed in Lejeusne *et al.* 2010), while in the Baltic, predicted higher temperatures and increased runoff may lead to lower salinity, eutrophication and increased hypoxia events (Andersson *et al.* 2015). Future responses to climate change are expected to differ both across and within species (Hare *et al.* 2016; Kjesbu *et al.* 2022), raising the question of what the differences in climate change vulnerability across a species' range might be.

Anadromous sea lamprey *Petromyzon marinus* are native to both sides of the North Atlantic Ocean, distributed from the Labrador Sea to the Gulf of Mexico in the west, from the Barents Sea to North Africa in the east and the Mediterranean Sea (Potter *et al.* 2015). Sea lamprey larvae live in fresh water until metamorphosis, followed by a marine phase spent parasitizing other fish, ultimately undertaking a spawning migration into rivers after which they die (Hardisty & Potter 1971; Renaud & Cochran 2019). Threat levels vary across the sea lamprey native range, with many locations considered of conservation concern (NatureServe 2024; OSPAR status assessment; BDC 2022/sea lamprey *Petromyzon marinus*). Sea lamprey is assessed as Least Concern globally by the IUCN Red List (Ford 2024). Meanwhile at the local level, many locations are threatened while others are not. Mismatch between species- and local-level assessments strengthens the argument for a more localized approach to sea lamprey conservation. Predicting future trends and understanding the differences in threats experienced

across the sea lamprey native range will also be valuable as future range shifts are expected due to rising temperatures affecting sea lamprey physiology and phenology (Wang *et al.* 2021).

Marine biodiversity loss and shifts in distribution due to changing environmental conditions and human activity are already underway and expected to intensify (Sala & Knowlton 2006). Anadromous fishes are particularly vulnerable to environmental changes; shifts in temperature and flow regimes can lead to range shifts, population declines and extirpations, and changes in life history (Lassalle & Rochard 2009; Limburg & Waldman 2009; Waldman & Quinn 2022). Adaptation is one way species can respond to impending climate change. Widely distributed species inhabiting heterogeneous environments experience spatially varying selection pressures that can drive local adaptation (Kawecki & Ebert 2004). Uncovering adaptation to environmental gradients across a species' range and locality-specific differences in life history has the potential to assist in predicting species- and population-level responses to future changing conditions. While conservation initiatives benefit from information on current adaptation, incorporating predictions of the strength of future selection pressures expected across a species' range is also valuable. Adaptation is not a static process; understanding dynamics across time and accounting for different climate change scenarios could help forecast future adaptive ability and tailor conservation planning with long-term environmental changes in mind. I therefore conduct an initial scan for adaptation across natural and human-impacted environments, followed by estimating genome-environment mismatch under future climate change.

Signals of adaptation have been detected in Pacific lamprey *Entosphenus tridentatus*, another anadromous parasitic lamprey species with a similar life-history. In the northeastern Pacific part of their range, outlier loci have been connected to geography, whether run timing

occurred earlier or later in the season, and dwarf life history, a smaller body form of Pacific lamprey (Hess *et al.* 2013). Local adaptation to individual body size and maturity traits have also been detected, indicating potential spatiotemporal differences in fitness connected to body size and ocean maturing vs stream maturing ecotypes (Hess *et al.* 2020). Adaptation linked to environmental conditions and life history could be occurring in anadromous sea lamprey across the Atlantic, with no studies to date spanning their entire distribution range. Meanwhile, sea lamprey in the East Atlantic display spatial selection connected to oxygen concentration and river runoff, metrics putatively related to environment and human-caused pressures (Baltazar-Soares *et al.* 2023). Abundance of European hake *Merluccius merluccius* and Atlantic cod *Gadus morhua*, sea lamprey host species during their parasitic marine phase, could also be a driver of selective pressure, but the authors were unable to uncover the exact mechanisms underlying this relationship (Baltazar-Soares *et al.* 2023).

Sea lamprey from across the North Atlantic occupy a variety of different habitats, with fluctuations across space potentially causing local adaptation. The limits placed on sea lamprey distribution by temperature are not known, but they reportedly inhabit temperatures between –0.6 and 20^{o}C (reviewed in Beamish 1980), and oxygen consumption is known to rise with temperature (Beamish 1973). In the upper ocean, oxygen concentrations increase with latitude while temperature decreases (Deutsch *et al.* 2020). Temperature and oxygen, highly correlated marine environmental variables, are connected to ectotherm thermal regulation (Pörtner & Knust 2007), metabolic processes (Deutsch *et al.* 2020, 2024; Pörtner & Knust 2007) and habitat breadth (Deutsch *et al.* 2020, 2024) in the marine environment. The study of temperature and oxygen as drivers of adaptation could also help predict future responses to rising temperatures and ocean deoxygenation conditions.

Environmental pH contributes to acid–base compensation in fish through metabolic processes facilitating acid–base transfer between organism and environment (Gilmour & Perry 2009). Metabolic requirements thus differ across environmental gradients and could drive local adaptation. Coastal systems are complex and influenced by watershed processes, nutrient inputs and ecosystem metabolism causing changes in water pH (Duarte *et al*. 2013; Hendriks *et al.* 2015). Adaptation to pH gradients sea lamprey encounter across their range could help understand response to future ocean acidification, with pH expected to decrease by an average of up to around 0.33 ± 0.04 globally by the end of the century (Jiang *et al.* 2019).

Salinity is known to drive adaptation across environmental gradients in fish species (Berg *et al.* 2015; Gaggiotti *et al.* 2009; Guo *et al.* 2016; Riccioni *et al.* 2013). Individuals from Sweden that inhabit a transitional zone but have not been able to establish in the lower salinity waters of the Baltic Sea (Mikael Svensson, SLU Swedish Species Information Centre, pers. comm.) are of specific interest in this study. Swedish sea lamprey inhabit the Baltic to North Sea transition zone, an area with a strong salinity gradient where the high salinity North Sea meets the low salinity Baltic Sea (Maar *et al.* 2011). A widely distributed species, sea lamprey encounter diverse environmental conditions across their range and adaptation to local salinity levels could potentially limit sea lamprey dispersal at sea. The salinity range sea lamprey are exposed to during their marine feeding phase is unknown. However, individuals used in this study inhabit environments with a salinity (PSS) range of 20.25 (Sweden) to 37.71 (French Mediterranean coast) (data extracted as part of this study from Bio-ORACLE; Assis *et al.* 2018; Tyberghein *et al.* 2012).

Sea lamprey may also experience selection pressure to human-impacted environments. As the human population continues to increase (United Nations Department of Economic and

Social Affairs, Population Division 2022) so will its evolutionary impact (Steffen *et al.* 2011). As an anadromous species, sea lamprey are vulnerable to the effects of human activity, especially during their migratory phase. Barriers to migration and habitat degradation (Maitland *et al.* 2015) could be forcing sea lamprey to adapt as many of their migration routes traverse highly populated, urbanized areas. Human activity is a known driver of species evolution (Johnson & Munshi-South 2017; Palumbi 2001), a relationship also reported in Chapter 2 of my thesis. It could be possible that difficulties navigating human-impacted migration routes and selective pressures due to habitat degradation factors such as pollutants, habitat alteration, and predation at migration barriers near large human population centers are causing local adaptation.

Sea lamprey have been increasingly observed in the northernmost parts of their range (Pereira *et al.* 2012; Van Leeuwen *et al.* 2021), and it is important to think about what future implications may be as climate change drives further northward range expansions. One perspective is that northward range expansion may have positive outcomes, compensating for loss of suitable habitat in the south (Lassalle *et al.* 2008). Another is that since sea lamprey are highly invasive and destructive in the Laurentian Great Lakes, northward movement due to climate change may open new environments for invasion. It would be sensible to consider the potential for freshwater adaptation as climate change increases habitat availability at higher latitudes. To date, Europe lacks any known landlocked populations, though feeding in fresh water before outmigration to sea has been recorded (Docker & Potter 2019). Individuals from Ireland sampled from a location where freshwater feeding is known to occur were found to be genetically distinct in Chapter 3, highlighting the need to understand adaptation in regard to freshwater residency. Although I have a limited number of samples, it is worth exploring this question to encourage future, more in depth, research efforts. Anadromy is thought to be

facultative and dependent on environment; sea lamprey need access to large water bodies with a sufficient prey base to become established in fresh water (Docker & Potter 2019). Therefore, managers may need to identify and monitor candidate freshwater environments for invasion as northern habitat becomes accessible under expected range shifts driven by changes in temperature and precipitation (Lassalle *et al.* 2008). Uncovering potential adaptation to freshwater environments will help assess anadromous sea lamprey's invasive potential in environments other than the Laurentian Great Lakes.

Sea lamprey, a highly dispersive species with a wide distributional range encompassing diverse environmental conditions, is an ideal candidate for investigating adaptation and selection. The strength of this study lies in sampling from across the entirety of oceanographic conditions anadromous sea lamprey experience across their distribution range and acquiring individuals from locations known for unique life history traits. My dataset consists of 155 individuals, including individuals from Sweden that are thought to avoid the lower salinity waters of the Baltic Sea, individuals from Ireland sampled where freshwater feeding is known to occur, and understudied Mediterranean sea lamprey. I plan on using two complementary approaches. First, I will test whether sea lamprey are locally adapted across environmental gradients and differences in human impact levels. I will then use this information to estimate genetic offsets, a metric of genome-environment mismatch under future climate change. Using genetic offsets, I will make predictions on the strength of selection pressure expected under future environments across the sea lamprey distribution range. These results will provide information for prioritizing future conservation efforts and assessing anadromous sea lamprey vulnerability per locality. I expect sea lamprey to be adapted to local environmental conditions and human-derived pressures. I also expect to find signs of adaptation to freshwater in Ireland, despite the low number of individuals

sampled. I predict that at more northern latitudes, the Mediterranean Sea and the Baltic to North Sea transition zone, sea lamprey will experience stronger selection pressures in respect to future environments and will therefore be more vulnerable to climate change.

Methods

Samples

I used the same sample set as seen in Chapter 3: 155 individuals from the East Atlantic (n = 109), West Atlantic (n = 40) and Mediterranean Sea (n = 6) (Table 4.1).

Sequencing and data processing

I used the same sequencing data as in Chapter 3. DNA was extracted with the Qiagen DNeasy Blood and Tissue Kit and sent for PCR-free library construction and whole genome sequencing (TCAG; The Hospital for Sick Children, Toronto, Canada). Samples were sequenced on two Illumina NovaSeq 6000 runs with paired 150-bp reads. Anadromous sea lamprey samples were sequenced alongside freshwater resident sea lamprey from the Great Lakes as part of a larger scale study.

Data processing, SNP variant calling and filtering were undertaken by Jaanus Suurväli (University of Manitoba) as part of the larger scale study and summarized below. Demultiplexed raw sequence data was received in .fastq format. The program fastp (v0.23.1) (Chen *et al.* 2018) was used to remove adapters and quality trimming. Quality control was based on reports from FastQC (v0.11.9) (Andrew 2015) and fastp. Reads were then aligned to the sea lamprey germline assembly kPetMar1.pri (Timoshevskaya *et al.* 2023) using BWA mem (v0.7.17) (Li & Durbin 2009). Duplicate reads were located and removed with MarkDuplicates in Picard (v2.26.3) (Broad Institute 2019). A unique read group identifier was assigned per sample with AddOrReplaceReadGroups in Picard. Samtools (v1.15.1) (Danecek *et al.* 2021) was used to

remove duplicate reads, reads that could not be mapped as a pair, reads that failed quality checks or had a quality score below 20. Genetic variants were located using Platypus (v0.8.1) (WGS500 Consortium *et al.* 2014)) and output .vcf files merged using vcf-concat from vcftools (v0.1.16) (Danecek *et al.* 2011). Variants passing default filters set by Platypus met the criteria for downstream analyses.

Anadromous sea lamprey individuals were then extracted from the full file and sorted by location using bcftools view (bcftools v0.1.11.). Genetic sites left without polymorphisms after removal of samples not used in this study were excluded using SelectVariants in GATK v4.2.6.1 (McKenna *et al.* 2010). Sites missing data for > 20% across all individuals were removed using bcftools.

Local adaptation tests

To minimize the likelihood of false positives, I used three different approaches to scan across the genome for putatively adaptive loci. The package pcadapt (v4.3.5) (Privé *et al.* 2020) detects loci that are outliers while controlling for population structure with Principal Component Analysis (PCA). Calculation of pairwise F_{ST} per site between the most genetically differentiated populations in vcftools (v0.1.16) allows isolation of the most highly differentiated genetic variants which can be considered putatively adaptive loci. Lfmm2 from the package LEA (Frichot & François 2015) is a method that uses latent factor mixed models (LFMM) to detect local adaptation with genome-wide association tests with respect to environmental gradients while controlling for population structure. A .lfmm-formatted file produced using the vcf2lfmm function in LEA (Frichot & François 2015) was used as input for all three approaches. I then ran analyses in R v4.3.2 unless otherwise stated.

For the first approach, I used the package pcadapt (v4.3.5) to detect outliers while controlling for population structure. In Chapter 3, I found divergence to be primarily trans-Atlantic. Based on this finding I set k = 2 genetically distinct populations, one on the East and one on the West Atlantic. I then assigned each genetic site a p-value representing correlation significance between principal components and data. I corrected for multiple testing using the false discovery rate method from the qvalue (v2.34.0) package, and considered sites with a q value < 0.05 putatively adaptive.

For the second approach I calculated pairwise F_{ST} per site between East and West Atlantic sea lamprey, the populations accounting for most divergence, using --Weir-fst-pop from vcftools (v0.1.16). I considered sites with the highest differentiation, represented by values in the highest 1%, to be putatively adaptive loci.

For the third approach, I used the tools lfmm2 and lfmm2.test from the package LEA v3.14.0 (Frichot & François 2015) to determine adaptation with respect to the environment. The value for k, the number of genetically distinct populations, was set to k=2 based on neutral population structure results (Chapter 3) that detected trans-Atlantic divergence. I examined adaptation to current environmental conditions using mean dissolved molecular oxygen ($mmol/m^{-3}$), mean surface pH, mean surface salinity (PSS) and mean sea surface temperature (ºC) (Bio-ORACLE; Assis *et al.* 2018; Tyberghein *et al.* 2012). These metrics reflect natural oceanographic conditions that could drive evolution and are connected to climate change. To test for adaptation to human altered environments such as barriers to migration, habitat degradation and their subsequent effects on life history, I used human population density as a proxy. Estimates of human population density (number of persons per square kilometer) were based on information from national censuses and population registers, for the year 2010 in a 30 arc-second

resolution and accessed from CIESIN (2018). Mean values were extracted from these datasets using a 200 km radius buffer. I retained the original dataset coordinate systems (WGS84) as it was shared between datasets. In the case that metadata associated with anadromous sea lamprey did not include coordinates, or coordinates were in freshwater, I took marine coordinates from the drainage basin of the water body individuals were sampled from. In the case of Iceland, where sea lamprey are not known to spawn and individuals were caught opportunistically (mostly at sea), I averaged coordinates across individuals. To test for potential adaptation to freshwater residency, I used a binary variable where 1 represented individuals caught at a location in Ireland found to be genetically distinct in Chapter 3 where feeding in fresh water before outmigration is known to occur, and 0 represented individuals with typical marine feeding habits. I then assigned each genetic site a significance value and corrected for multiple testing using the package qvalue (v2.30.0) and a significance threshold of 0.05 to identify candidate adaptive loci of interest.

Gene ontology

To detect which genetic variants are associated with which genes, and which genes correspond to known gene ontology (GO) annotations, I used variant annotation and gene-GO association files produced as part of an ongoing study (Jaanus Suurväli, University of Manitoba, unpublished data). In summary, to produce the variant annotation file, a genome annotation for kPetMar1.pri in .gff format was downloaded from NCBI and tRNA information removed. Ensembl Variant Effect Predictor (VEP, v107.0) (McLaren et al. 2016) was run using the aforementioned file, a genome in .fasta format and genetic variants in the form of a thinned .vcf file as input. Genetic sites were considered associated with a gene when within 5 kb distance, and cases where a genetic variant was associated with two different genes while being within the

boundaries of one were addressed with a custom R script by removing the up or downstream gene annotation. Using the VEP annotation file, lists of candidate adaptive single nucleotide polymorphisms (SNPs) and their associated candidate genes could be produced in R. To produce the gene-GO association file, protein sequences associated with known kPetMar1.pri genes were downloaded from NCBI. Sequences were associated with GO terms by functional annotation using the Pannzer2 online server (http://ekhidna2.biocenter.helsinki.fi/sanspanz/ ; accessed 03 May 2022) (Törönen *et al.* 2018). I then merged GO annotations for isoforms encoded by protein-coding genes.

To minimize the likelihood of false positives, I considered four different approaches to subsetting candidate loci for subsequent GO analysis. For the first approach I detected statistically significant GO term/gene associations for variants associated with environmental gradients and human activity from lFMM2, using enrichment analysis with the enricher tool from clusterProfiler (v4.10.0) (Wu *et al.* 2021). I ran multiple testing corrections in enricher with a q-value significance threshold of 0.05. For the second approach, I detected GO term/gene associations for variants found across all three adaptation methods (LFMM2, pcadapt, F_{ST} outliers). For the third and fourth approach, I detected GO term/gene associations for candidate loci shared between LFMM2/pcadapt and LFMM2/F_{ST} outlier methods respectively. I used the online tool ReviGO v1.8.1 (http://revigo.irb.hr/ ; accessed 23 November 2023) (Supek *et al.* 2011) to interpret GO results and make inferences on the biological mechanisms underlying sea lamprey adaptation.

Allelic turnover & genetic offsets

I created an input .vcf file containing SNP variants found to be putatively adaptive by LFMM2, pcadapt and F_{ST} outlier methods during previous analysis and converted it to 012

format using vcftools. I note that for variants detected by LFMM2, only those found to be connected to oxygen, pH, salinity, temperature and human population density were included, as they are the variables under examination in the following analysis. I extracted metrics of mean surface dissolved molecular oxygen, mean surface pH, mean surface salinity, mean sea surface temperature (Bio-ORACLE; Assis *et al.* 2018; Tyberghein *et al.* 2012) and human population density (CIESIN 2018) in R using a 200 km radius buffer to use in this analysis. I consider these variables to be candidate underlying drivers of both contemporary and changing future environments. For this analysis I interpret human population density as an indicator of the effects of urbanization and associated sea lamprey habitat degradation, pollution and barriers to migration.

I used the method Gradient Forest (Láruson *et al.* 2022) as applied by the gradientForest R package (v0.1-37) (Ellis *et al.* 2012) to determine how "maladapted" sea lamprey potentially are to their respective future environments. In summary, this method uses a random forest approach to determine environmental associations with allele frequency turnover across space. It then estimates the genetic distance, or otherwise "mismatch" between present and predicted future environments at a location. I interpret future maladaptation risk as the strength of selection pressure expected to be experienced at a location. I ran the analysis on the entire dataset spanning the anadromous sea lamprey distribution range, an East Atlantic/Mediterranean subset, and a West Atlantic subset. I analyzed East and West subsets separately to explore coastal adaptive processes, as in Chapter 3 population divergence was primarily trans-Atlantic.

I initially accounted for spatial patterns using PCNMs (Principal Coordinates of Neighborhood Matrix), functions that use a truncation threshold for long distances and map between row distances onto a rectangular matrix on rows (Borcard & Legendre 2002), using the r

package vegan v2.6-4 (Oksanen *et al.* 2022). I calculated the association between genotype and environment across space using the gradientForest() function with ntree = 500, nbin = 301 as model parameters and SNPs, climate and spatial variables as input. Predictor variables and significant PCNMs were ranked by importance and plotted. I then estimated present day turnover of allele frequencies across environmental gradients using the predict() function and associated environmental variables and fitted gradient forest model as input. I then plotted the resulting map. Human population density was not included in allelic turnover calculation and only interpreted based on its predictor importance ranking in the model. It was included in the initial model as it is informative and worth consideration but lacks compatible mapped marine data. Allelic turnover calculation requires data that is constant across space while this metric of human population density consists of point data.

I then estimated turnover of allele frequencies across future environmental gradients with the predict() function, using the fitted gradient forest model and projections of future environments as input. Projected future environments were represented by metrics of forecasted mean surface salinity (PSS) and mean sea surface temperature (ºC) (Bio-ORACLE; Assis *et al.* 2018; Tyberghein *et al.* 2012) for the years 2040-2050 and 2090-2100, accounting for both shorter and longer-term effects. It is noted that forecasted datasets of future oxygen and pH levels are not available, so this part of the analysis focused on salinity and temperature. I considered salinity and temperature conditions for each of these two time periods for four respective RCPs (representative concentration pathway scenarios); RCP2.6, a peak and decline scenario predicting very low greenhouse gas levels by the end of the 21st century, RCP4.5 and RCP6.0 predicting stabilization of greenhouse gas levels and PCP8.5 that predicts emissions increase over time leading to high greenhouse gas levels. Running predictions for each period

under each likely climate change scenario will account for sea lamprey maladaptation risk in both closer and more distant future under best, stable and worst-case scenarios. I then calculated genetic offsets, the per location genetic distance between present and future environments, using Euclidean distance as a distance metric and estimations of allelic turnover across present and future environmental gradients as input. Genetic offsets are values representing the change in allele composition needed to maintain local adaptation under future environments (Láruson *et al.* 2022). I then mapped and plotted boxplots of genetic offsets by geographic sampling site.

Results

Local adaptation & gene ontology

By testing for adaptation across environmental gradients using lfmm2 in LEA (Frichot & François 2015), I detected 7587 SNP outliers associated with freshwater feeding in Ireland, 10 associated with sea surface temperature, 45 associated with surface salinity, 142 associated with dissolved oxygen, 3131 associated with surface pH and 198 associated with urbanization. Using Principal Component Analysis in pcadapt (Privé *et al.* 2020), I detected 7020 candidate outliers. I then identified 2947 outlier SNPs using the highest 1% differentiated sites based on pairwise F_{ST} per site. I found 25 SNP outliers associated with freshwater feeding in Ireland, 1 associated with surface salinity, 1 associated with dissolved oxygen and 60 associated with surface pH shared across all three outlier detection methods. I found no candidate adaptive loci for sea surface temperature or urbanization based on this approach.

The joint consideration of LFMM2 and pcadapt methods enabled me to locate 31 SNP outliers associated with freshwater feeding in Ireland, one associated with sea surface temperature, three associated with surface salinity, nine associated with dissolved oxygen, 228

associated with surface pH and 11 associated with urbanization. The joint consideration of LFMM2 and the F_{ST} outlier methods enabled me to locate 41 SNP outliers associated with freshwater feeding in Ireland, two associated with surface salinity, one associated with dissolved oxygen and 65 associated with surface pH. I found no adaptive loci associated with sea surface temperature or urbanization. Outlier SNPs detected across methods can be seen in Figure 4.1. and Figure 4.2.

I initially detected statistically significant GO terms for variants identified by LFMM2, on which I ran an enrichment analysis. I found variants linked to environmental gradients and human activity using LFMM2 to be associated with the following number of genes: 3429 for freshwater feeding in Ireland, 7 for sea surface temperature, 31 for surface salinity, 100 for dissolved oxygen, 1929 for surface pH and 150 for urbanization. Enrichment analysis detected the following GO terms: 172 significant GO terms for freshwater feeding in Ireland and 207 for surface pH. I found no significant GO terms for sea surface temperature, surface salinity, dissolved oxygen and urbanization.

I then examined variants overlapping all three methods used to detect adaptation (LFMM2, pcadapt, F_{ST} outliers). These variants were associated with the following number of genes: 15 for freshwater feeding in Ireland, 1 for surface salinity, 1 for dissolved oxygen and 47 for surface pH. These genes were connected to 433 GO terms for freshwater feeding in Ireland, 34 for dissolved oxygen and 2446 for surface pH. Surface salinity had no associated GO terms.

When considering variants shared between the LFMM2 and pcadapt methods, I found them to be associated with the following genes: 19 for freshwater feeding in Ireland, 1 for sea surface temperature, 2 for surface salinity, 8 for dissolved oxygen, 182 for surface pH and 8 for urbanization. These genes were connected to 457 GO terms for freshwater feeding in Ireland,

471 for dissolved oxygen, 10058 for surface pH and 580 for urbanization. Surface salinity and sea surface temperature had no associated GO terms.

Finally, I considered variants shared between the LFMM2 and F_{ST} outlier methods. I found these outlier loci to be associated with the following number of genes: 25 for freshwater feeding in Ireland, 2 for surface salinity, 1 for dissolved oxygen and 51 for surface pH. These genes were connected to 997 GO terms for freshwater feeding in Ireland, 34 for dissolved oxygen, and 2784 for surface pH. Surface salinity had no associated GO terms.

Freshwater feeding behavior in Ireland and pH are the most strongly supported drivers of adaptation. They were supported by all approaches used to detect adaptation; variants identified by LFMM2 and subsequent enrichment analysis, variants shared between LFMM2, pcadapt and F_{ST} outliers, variants shared between LFMM2 and pcadapt, variants shared between LFMM2 and F_{ST} outliers. Adaptation to dissolved oxygen was identified by three of the approaches used; variants shared between LFMM2, pcadapt and F_{ST} outliers, variants shared between LFMM2 and pcadapt, variants shared between LFMM2 and F_{ST} outliers. However, I found limited numbers of both candidate SNPs and associated genes; therefore, the significance of oxygen as a driver of adaptation is negligible. Lastly, variants shared between LFMM2 and pcadapt identified urbanization as a driver of adaptation, but I found limited numbers of both candidate SNPs and associated genes. As it was detected by only one out of four approaches and is associated with a very limited number of candidate SNPs and genes, urbanization is not a likely driver of adaptation. Overall, my results indicate that anadromous sea lamprey are adapted to their local environments.

Gene ontology interpretations for anadromous sea lamprey should be approached with caution. Most GO term interpretations are based on knowledge available for model organisms.

Since they are not closely related to sea lamprey, these results may not be translatable; for example, a gene connected to a known biological function in the house mouse *Mus musculus*, a mammal, may not maintain the same biological function in sea lamprey, a jawless fish. However, these results are still interesting to consider.

ReviGo interpretations of GO results for variants identified by LFMM2 and subsequent enrichment analysis, and variants shared between LFMM2, pcadapt and F_{ST} outliers, the more stringent subsets of adaptive loci detected, can be seen in Supplemental Information 3.1 and 3.2. I found no overlap in ReviGo interpretations between these two approaches for pH. Biological functions associated with GO terms varied and no clear pattern was observed. For freshwater feeding, overlap between ReviGo interpretations for variants identified by LFMM2 and subsequent enrichment analysis, and variants shared between LFMM2, pcadapt and F_{ST} outliers, was found for five GO terms. The biological functions implicated in these five overlapping GO terms are "secretion by cell", "cell junction organization", "synapse organization", "postsynapse organization" and "export from cell", and seem to be related to cellular transport and organization. These could be implicated in cellular processes taking place during osmoregulation (Ferreira-Martins *et al.* 2021), as sea lamprey acclimatize to either freshwater or marine environments during migration. Additionally, changes in the structure of junctions in the gill epithelium connected to osmoregulation have been observed during migration in anadromous lampreys (reviewed in Bartels & Potter 2004).

Allelic turnover & genetic offsets

Results for the entire dataset are presented in this section. Results for the East Atlantic/Mediterranean and West Atlantic subsets were consistent with and reflected those found in the entire dataset and can be seen in Supplemental Information 2 as noted below. PCNMs

calculated for the entire dataset showed that the most significant spatial patterns are primarily connected to trans-Atlantic divergence and a latitudinal gradient (Supplemental Information 2.16-2.18), consistent with population structure analysis results seen in Chapter 3. PCNMs calculated for the East Atlantic/Mediterranean and West Atlantic subsets can be seen in Supplemental Information 2.19-2.22 and Supplemental Information 2.23-2.24, respectively, further confirming that spatial patterns are predominantly latitudinal. Fitted model predictor variables and significant PCNMs for the entire dataset were ranked by their importance in the model (Figure 4.3). Equivalent plots for the East Atlantic/Mediterranean and West Atlantic subsets are presented in Supplemental Information 2.25, 2.26 respectively. PCNM1 (Supplemental Information 2.16), representing trans-Atlantic divergence, followed by PCNM2 (Supplemental Information 2.17), representing the latitudinal gradient, were found to best explain variation in the model. Therefore, genotype associations across the anadromous sea lamprey native range are primarily spatial.

Surface salinity, followed by surface pH and, to a lesser extent, human population density as a metric of pollution and urbanization were found to be the next most important predictor variables in the model. Surface salinity and surface pH are the most significant genotype/environment associations, despite spatial effects being the primary driver. Human population density was less significant. PCNM3 (Supplemental Information 2.18) was found to be the least important spatial predictor while surface dissolved oxygen and sea surface temperature were found to be the least important environmental predictors in the model.

Present day turnover of allele frequencies across space and environments, interpreted as continuous changes in inferred allele frequencies associated with candidate environmental predictors can be seen in Figure 4.4 for the entire dataset and in Supplemental Information 2.27,

1.28 for the East Atlantic/Mediterranean and West Atlantic subsets, respectively. Increasingly differentiated rates of allelic turnover are observed at higher latitudes and the Baltic Sea, environments highly affected by climate change (Andersson *et al.* 2015; Bonebrake & Mastrandrea 2010; Screen 2014), reflecting more pronounced local environmental differences in salinity, pH, oxygen and temperature and subsequent local adaptation.

Genetic offsets, the forecasted likelihood of maladaptation, are here represented by the predicted mismatch between present and predicted future allelic turnover. Predictions of maladaptation to respective future environments remained consistent across RCP scenarios and two alternate timescales, and plots of climate change scenarios and timescales not presented here can be found in Supplemental Information 2.29-2.42 for the entire dataset, in Supplemental Information 2.43-2.58 for the East Atlantic and Mediterranean region, and Supplemental Information 2.59-2.74 for the West Atlantic. Mapped genetic offsets across space and gradients in future salinity and temperature for a RCP6.0 climate change scenario for the years 2090-2100 across the entire sea lamprey native range are presented in Figure 4.5 while boxplots of the same genetic offsets plotted per site can be seen in Figure 4.6. Sea lamprey from Sweden and New Brunswick were found to be the most at risk of maladaptation. Sea lamprey from Iceland and Germany were the next most maladaptive, though noticeably less vulnerable than those in Sweden and New Brunswick. Individuals from the French Atlantic coast (rivers Oir, Loire and Scorff), Ireland (rivers Moy and Shannon), rivers Mondego, Tejo and Vouga in Portugal and Massachusetts were found to be intermediately vulnerable while individuals from the French Mediterranean Coast, the River Guadiana in Portugal, North Carolina (rivers Roanoke, Pamlico/Tar, Neuse, Pee Dee) and Maryland were found to be the most resilient.

Discussion

My initial scan across the sea lamprey genome for adaptive loci uncovered that sea lamprey are indeed adapted to their local environments. I found adaptation across a pH gradient and adaptation to freshwater feeding behavior to be the most strongly supported. I then found that risk of maladaptation follows a latitudinal gradient. Sweden and New Brunswick, two of the higher latitude locations are most vulnerable under future conditions while locations at lower latitudes are the least vulnerable.

Sea lamprey seem to be adapted to their local environments, with freshwater feeding behavior in Ireland and pH being the strongest candidate drivers of adaptation. Adaptation to freshwater feeding behavior should be taken with caution as this study does not include enough individuals to provide conclusive results. Feeding in fresh water prior to out-migration is known to occur at multiple locations throughout the anadromous sea lamprey native range (reviewed in Docker & Potter 2019). In North America instances have been recorded in the St. John River in New Brunswick (Beamish 1980) and more often in Love Lake in Maine (Davis 1967). In Europe, freshwater feeding has been observed in the Rhine River in Germany (Baer *et. al* 2018) and Loch Lomond in Scotland (Maitland 1980; Maitland *et al.* 1994), and with higher frequency in the River Ulla and River Umia in Spain (Silva et al. 2013b, a). Most habitual freshwater feeding behavior has been observed in multiple Irish lakes (King & O'Gorman 2018). In my study a total of 9 individuals were sourced from Lough Conn in Ireland, a location at which freshwater feeding has been observed, but is not the norm. However, these individuals were sampled haphazardly when ideally individuals found feeding parasitically in fresh water would have been targeted. While it is unknown whether individuals sourced from Lough Conn fed while freshwater resident, I found 4 individuals to be genetically different during population

structure analysis in Chapter 3, but they are unfortunately too few to base analysis upon. Freshwater feeding behavior has also been observed in Lough Derg, another location used in this study. Inclusion of these additional 8 individuals, unknown whether they have exhibited freshwater feeding behavior, would still not be sufficient to reach conclusive results. However, my findings in both this chapter and Chapter 3 encourage future research on the potential for sea lamprey adaptation to freshwater environments other than the Laurentian Great Lakes. I would recommend the genetically distinct Lough Conn in Ireland as a study system, using either a larger sample size or, if possible, targeting individuals known to have fed in fresh water. Results from this proposed study could then be compared with genes found to be under selection for freshwater residency in the Laurentian Great Lakes (Jaanus Suurväli, University of Manitoba, unpublished data).

I consider variation in pH the main driver of adaptation in anadromous sea lamprey across the North Atlantic and Mediterranean Sea. Individuals used in this study inhabit environments fluctuating from pH 7.87 (New Brunswick) to 8.27 (North Carolina) (data extracted as part of this study from Bio-ORACLE; Assis *et al.* 2018; Tyberghein *et al.* 2012) across the sea lamprey native range. Acid–base compensation is mediated by acid–base transfer between fishes and their respective environments through metabolic processes (Gilmour & Perry 2009). Energetically intense spawning migrations from marine to fresh water and the need to maintain homeostasis across drastically different environments could place high energy demands on anadromous lampreys. In certain lamprey species, pH is associated with metabolic processes required to meet high energy requirements (Nikinmaa 1993; Nikinmaa & Mattsoff 1992; Tufts *et al.* 1992). These cellular mechanisms could be underlying local adaptation across environmental gradients in pH, locality-specific pH levels connected to differences in metabolic and energy

requirements. Spatial variation in selective pressure could explain outlier loci connected to differences in pH levels and be driving limits to sea lamprey dispersal. Water pH has been identified as a selective pressure for both marine and freshwater fishes across different habitats (Delaval *et al.* 2022; Haenel *et al.* 2019; Hay *et al.* 2022), though studies are limited in comparison to other abiotic parameters. Baltazar-Soares *et al.* (2023) proposed that local adaptation to abiotic conditions in East Atlantic anadromous sea lamprey is likely driven by spatially varying selection, though pH was not included in this study. Spatially varying selection has also been detected in American eel *Anguilla rostrata* (Gagnaire *et al.* 2012), another North Atlantic diadromous fish, and similar evolutionary forces may be influencing anadromous sea lamprey. Meanwhile, ocean acidification poses a significant threat to marine organisms as pH regulation requires significant metabolic energy (Raven *et al.* 2005). Teleost fishes and other organisms with high metabolic rates seem to be less sensitive to acidification, potentially due to internal pH regulation linked adaptation (Melzner *et al.* 2009). Signals of adaptation to different pH levels across the anadromous sea lamprey distribution range encourage hopes of adaptive ability in the face of climate change. However, changes in ocean pH are already underway and could contribute to future climate change driven range shifts with unknown consequences.

Spatial variation in sea lamprey size is thought to occur across their native range. In the East Atlantic, sea lamprey display latitudinal variation in size, with the largest individuals found in Portugal and size declining from south to north (Beaulaton *et al.* 2008). In the West Atlantic, sea lamprey are smaller than in the East Atlantic and seem to follow an opposite latitudinal trend, with largest individuals found in the North and smallest in the South (reviewed in Beaulaton *et al.* 2008). No reports of sea lamprey size are published for the Mediterranean Sea. However, observed differences could be biased by sea lamprey size decreasing as the spawning migration

progresses, a trait that differs in magnitude between males and females (Docker & Potter 2019). Adaptive markers connected to differences in body size and spawning migration distance have been found in anadromous Pacific lamprey (Hess *et al.* 2014), and body size is likely connected to upstream migration distance (Hess et al. 2014; Spice et al. 2012). As sea lamprey size varies spatially, although not as pronounced as Pacific lamprey (Docker & Potter 2019), these studies raise the question of if sea lamprey size differences are also adaptive in nature and connected to upstream migration distance. While some of my samples were acquired through collaborators specifically for the purpose of my study, I also took advantage of archived samples acquired as part of different studies and museum specimens. Sampling efforts were also limited by logistics; for example, in some cases it was easier for collaborators to collect larvae. The tradeoff of my dataset having broad geographical coverage is that it does not include either adults for all locations or size measurements for all adults, and where adult size was included, it was not necessarily measured at the same time in their life cycle. For example, my dataset includes individuals sampled opportunistically during their marine phase and individuals sampled during their upstream migration. Additionally, it is uncertain if all spawning phase adults sampled as part of my study were captured at the same stage of their spawning migration to prevent bias due to size declines as migration progresses (Docker & Potter 2019). However, future studies could be planned to sample upstream migrating sea lamprey at the same migration stage (e.g., as they enter fresh water or just prior to spawning), across all locations, to shed light on these questions.

Predictions of future sea lamprey vulnerability are based on a RCP6.0 climate change scenario for the years 2090-2100 across their entire native range. This accounts for a longer time scale and a moderate climate change scenario that predicts stabilization of greenhouse gas levels by the end of the 21st century. By focusing on a moderate climate change scenario, I aim to

neither underestimate the threat of climate change nor be alarmist in my science. I found sea lamprey from Sweden and New Brunswick, two of the most northern locations sampled, to be the most likely to be maladapted under forecasted climate change. Sampling locations in Sweden are situated at the North Sea to Baltic transition zone, an area where two seas with very different biotic and abiotic processes, the continental shelf North Sea connected to the Atlantic and the enclosed brackish Baltic Sea (reviewed in Maar *et al.* 2011), meet. In the Northern Hemisphere (not including the tropics), higher latitudes are expected to be vulnerable to climate change (Bonebrake & Mastrandrea 2010; Screen 2014), and Swedish waters are likely influenced by the adjacent Baltic Sea, a climate change hotspot (Andersson *et al.* 2015). I expect these locations will be most vulnerable to future environmental changes, with sea lamprey in these regions facing more intense selection pressure under future changes in temperature and salinity. My initial research plans included sampling efforts in Newfoundland, the most northern known sea lamprey spawning location. These plans were cancelled due to the COVID-19 pandemic. Given the vulnerability I detected at the more northern parts of the sea lamprey range it is unfortunate that individuals from Newfoundland were excluded from my study, and their inclusion would be an interesting future research direction. Genetic offset results from Iceland should be interpreted with caution as no breeding has been observed at this location. However, sea lamprey from northern locations are more likely to stray and be caught in Icelandic waters, thus vulnerability in Iceland can be considered an additional indicator of vulnerability in northern locations.

I found individuals from the French Mediterranean coast, the River Guadiana in Portugal, North Carolina and Maryland least likely to be maladapted, vulnerability seemingly following a latitudinal gradient. I expect sea lamprey from more northern locations will be mostly influenced by increases in temperature while sea lamprey from Sweden inhabiting the North to Baltic Sea

transitional zone will be under the joint influence of increasing temperatures and decreasing salinity. While sea lamprey in New Brunswick are considered Secure (S5; NatureServe 2024), Swedish sea lamprey are assessed as Endangered (BDC 2022/sea lamprey *Petromyzon marinus*). Despite expectations of climate change vulnerability in Mediterranean waters, I didn't find sea lamprey to be at significant risk of maladaptation at this location. Vulnerability of the Mediterranean basin to climate change (Cos *et al.* 2022; Diffenbaugh *et al*. 2007) seems to not be translating to the genetic level, possibly because strong selection pressures due to climate change have already driven change in the region. Lower maladaptation at the southernmost part of the sea lamprey distribution range could support contribution of rear edge populations to adaptive capability under climate change (Habibzadeh *et al.* 2021; Rehm *et al.* 2015). I expect sea lamprey in North Carolina, Maryland and Portugal to be best adapted in the face of impending environmental changes. Resilience at these locations should be higher at least in the short –term, and they are likely to contribute beneficial alleles through gene flow to less well adapted locations. Their potential for resilience and genetic rescue supports the importance of their conservation. However, climate change driven range contractions are expected in the southernmost parts of the sea lamprey distribution range (Lassalle *et al.* 2008), and these locations have likely already been subject to more intense selection pressures, raising questions about long-term survivability.

My results could serve as a warning to make efforts for the conservation of sea lamprey at locations expected to fare worse under climate change. These are generally expected to be the southern parts of the sea lamprey native range, while changes in host abundances and effects of changing biotic conditions on life history will also determine future distribution (reviewed in Maitland *et al.* 2015). My results could also be used to direct conservation efforts to populations

with both a higher likelihood of survival under current conditions and higher adaptive potential. Better adapted, highly resilient populations could also be a source of genetic rescue under future environmental conditions that lead to declines in those less well adapted. However, it would be important to initially take into consideration underlying population structure and match recipient locations with genetically similar source populations. As I found that sea lamprey are adapted to their local environments, it would also be sensible to consider translocations between locations with more similar environmental conditions. High levels of gene flow in sea lamprey (see Chapter 3) could enable the spread of beneficial alleles within populations on each Atlantic coast. High standing genetic variation drives adaptive ability (Barrett & Schluter 2008), with distinct locations benefiting from a high effective population size. I therefore interpret maladaptation likelihood in regards to effective population size (Ne) estimates calculated as part of Chapter 3 (Table 4.2) and whether sea lamprey are of conservation concern for each location.

Sea lamprey in New Brunswick are one of the locations most likely to be maladapted in the future. Sea lamprey are considered Secure in New Brunswick based on assessment by the Atlantic Canada Conservation Data Center (http://accdc.com/index.html), (S5; NatureServe 2024) and therefore not likely a conservation priority. An effective population size estimate is not available, but I consider this an artifact of large population size. Detected future vulnerability at this location is unsurprising as northern latitudes become increasingly susceptible to climate change. These results show that present day stable status does not necessarily predict future trends. The assumption that lamprey would fare better in a region where they are currently more abundant does not apply in this case, though high population numbers may be able to temporarily buffer against large scale declines. My results show that large population sizes are not a

guarantee of future resilience, and sea lamprey at this location should be carefully monitored to enable their persistence.

Sea lamprey in Sweden are considered Endangered based on their OSPAR status assessment (BDC 2022/sea lamprey *Petromyzon marinus*). Based on my results, this location is particularly vulnerable to future climate change, which is concerning, and low effective population size at this location adds to these concerns. My results encourage the protection of sea lamprey in Sweden and provide evidence that could help convince overseeing authorities to support conservation efforts.

Sea lamprey do not appear to be at risk of maladaptation in Portugal, encouraging hopes that a sustainable regional fishery could be supported in the future. Sea lamprey in Massachusetts, Portugal, Ireland and France, locations I found to be intermediately maladapted, have a high effective population size. The combination of a high effective population size and a moderate chance to be maladapted to future environmental conditions indicates potential for sea lamprey survivability at these locations.

Sea lamprey in the Mediterranean, North Carolina and Maryland are some of the least likely to be maladapted but have low effective population size. Despite these locations being more likely to be adapted to future environments, low effective population size could prevent survival and lead to loss of adaptive potential. In Germany, I found sea lamprey to be moderately maladaptive but with a low effective population size, meaning persistence at this location is likely also low.

It is also important to consider the limitations of this study. I used projections of future temperature and salinity due to limited data availability for future environments. There are likely many other environmental parameters driving future changes, such as changes in precipitation

and flow regimes, host abundance, and higher productivity in fresh water compared to marine habitats potentially offsetting the benefits of an anadromous life history (reviewed in Maitland *et al.* 2015). I also expect future environments to be unpredictable due to climate change changing known weather patterns and climatic processes, and sea lamprey to undergo range shifts; thus, predictions based on current knowledge may not necessarily be applicable in the future.

Genetic rescue, the translocation of individuals aiming to strengthen genetic diversity and adaptive potential, has been successful in restoring threatened populations (Novak *et al.* 2021). Adult Pacific lamprey undertaking their spawning migration are translocated from the lower Columbia River basin upstream to locations where they are locally extirpated or in decline, to increase abundance while efforts for habitat restoration and barrier passage are underway (Hess *et al.* 2022, 2023; Ward *et al.* 2012). These long-term efforts have shown to be successful, enhancing abundance across all phases of the Pacific lamprey life cycle (Hess *et al.* 2022, 2023). Although these translocations took place within the same river system, their success encourages considering this practice for sea lamprey. Translocated individuals must be able to survive in their new environments, and it is vital that local adaptation is considered when planning for genetic rescue. Combining knowledge on population vulnerability in regards to respective environment under climate change, population structure and status, could identify populations that have the potential to be both a source of genetic rescue and future resilience. Since my results show that sea lamprey are adapted to their local environments, individuals may need to be translocated between locations with similar environmental conditions. I would recommend further research on local adaptation dynamics between candidate translocation sources and recipient locations, and incorporating more environmental metrics.

In Sweden, restocking is being considered to mitigate sea lamprey declines in the region (Mikael Svensson, SLU Swedish Species Information Centre, pers. comm.). Since I found Swedish sea lamprey to be part of a larger, admixed population on the East Atlantic coast in Chapter 3, population genetic structure results permit multiple options for translocation sources. Translocations from Ireland, France and Portugal are all viable options based on sea lamprey status at these locations. As Portuguese sea lamprey are the most likely to be best adapted under future conditions, they could be a good translocation source. Sweden is part of the large, admixed East Atlantic population, and sea lamprey lack natal homing behavior (Bergstedt & Seelye 1995); therefore, immigrants may be occurring from multiple locations. Out of these, Ireland is the most likely source based on status and geographical distance. France and Portugal may not be closely located but could be a source of immigration based on their status. Germany would be a likely source as it is located close to Sweden but is unlikely to be providing many immigrants as lamprey are faring poorly in the region. Sea lamprey are also known to spawn in Norway and Denmark, countries that are closer to Sweden. Status in these countries is Data Deficient and Near Threatened, respectively (OSPAR status assessment; BDC 2022/sea lamprey *Petromyzon marinus*). While they could be a source of immigrants to Sweden, they were not included in my study; therefore, presence and rates of gene flow are unknown. Potential immigrants to Sweden face lack of spawning and larval habitat due to barriers to migration and interference of hydroelectric infrastructure on river hydrodynamics (OSPAR status assessment; BDC 2022/sea lamprey *Petromyzon marinus*), highlighting the importance of habitat restoration as part of conservation efforts.

To date, Pacific lamprey translocations in the Columbia River have focused on adults (Hess *et al.* 2022, 2023; Ward *et al.* 2012). Individuals used for translocations would likely not

have made it past the barriers they were captured at and are translocated within the same river basin; therefore, their removal does not affect the overall population. If the same practice is to be considered for sea lamprey, the advice of local managers should be sought as to whether the local population can support removal of individuals to support a translocation program, as individuals would be removed from one river system to supplement another. Another option would be to restock candidate freshwater habitats with larvae reared in captivity, limiting removal of individuals from wild populations. Although there are still knowledge gaps on rearing larval sea lamprey in captivity, previous studies have shown it is feasible and techniques are transferable across species (Hume *et al.* 2024; Lampman *et al.* 2021; Moser *et al.* 2019), encouraging future studies. Concerns on whether lack of natal homing in sea lamprey (Bergstedt & Seelye 1995) will prevent the success of restocking efforts will also need to be addressed before implementing this practice. Since sea lamprey do not home to their natal streams, it is possible that when restocked individuals return to fresh water to spawn, instead of returning to the intended location, they disperse haphazardly. Despite concerns for Pacific lamprey, successful translocation programs observed a substantial number of returning adults, suggesting that despite lack of natal homing in this species spawning adults seem to display a wider-scale preference for their basin of origin (Hess *et al.* 2023). Sea lamprey are typically found in deep waters more than 350 km offshore (Beamish 1980; Hardisty 1986) and disperse more widely than Pacific lamprey (reviewed in Spice *et al.* 2012), potentially increasing chances individuals stray from their intended destination. Thus, future research on the rates of adult return to restocked locations is important prior to attempting wide-scale restocking.

Conservation priorities and resource allocation are largely a cultural issue. My study has identified locations at which sea lamprey will likely be vulnerable in the future, and where they

are predicted to be most resilient. The interplay between knowledge on current status and future predictions is valuable for managers as it can warn about unexpected threats, such as future vulnerability found for the New Brunswick area and likely nearby locations with similar environmental conditions. Decisions on whether to prioritize vulnerable locations or dedicate resources to protecting those with highest chances of survival will be up to local managers and public opinion. My results provide a scientific background to assist decision making, and conservation will be ultimately influenced by local attitudes and value placed on anadromous sea lamprey.

Acknowledgements

I would like to thank Ted Castro-Santos (U.S. Geological Survey), Mike Wilkie (Wilfrid Laurier University), Guillaume Evanno (INRA), Magnús Jóhannsson (MFRI), Benóný Jónsson (MFRI), Jan Baer (LAZBW), Fiona Bracken (University College Dublin), Catarina Mateus (Universidade de Évora), Thomas Evans (St. Mary's College of Maryland), Michael Fisk (NC Wildlife Resources Commission), Jeremy McCargo (NC Wildlife Resources Commission), Gabriela M. Hogue (North Carolina Museum of Natural Sciences), Mikael Svensson (SLU), Elisabeth Thysell (County Administrative Board in Halland) for the anadromous sea lamprey samples. I would like to thank Arfa Khan, Jessie L. Ogden and Phil Grayson for the technical support as part of this project. This research was enabled in part by support provided by the Prairies DRI and the Digital Research Alliance of Canada (alliancecan.ca).

References

Andersson, A., Meier, H.E.M., Ripszam, M., Rowe, O., Wikner, J., Haglund, P., et al. (2015). Projected future climate change and Baltic Sea ecosystem management. AMBIO, 44, 345–356.

Andrew, S. (2015). FastQC: A quality control tool for high throughput sequence data [Online]. http://www.bioinformatics.babraham.ac.uk/projects/fastqc/

Assis, J., Tyberghein, L., Bosch, S., Verbruggen, H., Serrão, E.A. & De Clerck, O. (2018). Bio-ORACLE v2.0: Extending marine data layers for bioclimatic modelling. Glob. Ecol. Biogeogr., 27, 277–284.

Baer, J., Hartmann, F. & Brinker, A. (2018). Abiotic triggers for sea and river lamprey spawning migration and juvenile outmigration in the River Rhine, Germany. Ecol. Freshw. Fish, 27, 988–998.

Baltazar-Soares, M., Britton, J.R., Pinder, A., Harrison, A.J., Nunn, A.D., Quintella, B.R., et al. (2023). Seascape genomics reveals limited dispersal and suggests spatially varying selection among European populations of sea lamprey (*Petromyzon marinus*). Evol. Appl., 16, 1169–1183.

Barrett, R. & Schluter, D. (2008). Adaptation from standing genetic variation. Trends Ecol. Evol., 23, 38–44.

Bartels, H. & Potter, I.C. (2004). Cellular composition and ultrastructure of the gill epithelium of larval and adult lampreys. J. Exp. Biol., 207, 3447–3462.

Beamish, F.W.H. (1973). Oxygen consumption of adult *Petromyzon marinu*s in relation to body weight and temperature. J. Fish. Res. Board Can., 30, 1367–1370.

Beamish, F.W.H. (1980). Biology of the North American anadromous sea lamprey, *Petromyzon marinus*. Can. J. Fish. Aquat. Sci., 37, 1924–1943.

Beaulaton, L., Taverny, C. & Castelnaud, G. (2008). Fishing, abundance and life history traits of the anadromous sea lamprey (*Petromyzon marinus*) in Europe. Fish. Res., 92, 90–101.

Berg, P.R., Jentoft, S., Star, B., Ring, K.H., Knutsen, H., Lien, S., et al. (2015). Adaptation to low salinity promotes genomic divergence in Atlantic cod (*Gadus morhua* L.). Genome Biol. Evol., 7, 1644–1663.

Bergstedt, R.A. & Seelye, J.G. (1995). Evidence for lack of homing by sea lampreys. Trans. Am. Fish. Soc., 124, 235–239.

Bonebrake, T.C. & Mastrandrea, M.D. (2010). Tolerance adaptation and precipitation changes complicate latitudinal patterns of climate change impacts. Proc. Natl. Acad. Sci., 107, 12581–12586.

Borcard, D. & Legendre, P. (2002). All-scale spatial analysis of ecological data by means of principal coordinates of neighbour matrices. Ecol. Model., 153, 51–68.

Broad Institute. (2019). Picard toolkit. https://broadinstitute.github.io/picard/

Center for International Earth Science Information Network (CIESIN) Columbia University. (2018). Gridded population of the world, Version 4 (GPWv4): population density, Revision 11.

Chen, S., Zhou, Y., Chen, Y. & Gu, J. (2018). fastp: An ultra-fast all-in-one FASTQ preprocessor. Bioinformatics, 34, i884–i890.

Cos, J., Doblas-Reyes, F., Jury, M., Marcos, R., Bretonnière, P.-A. & Samsó, M. (2022). The Mediterranean climate change hotspot in the CMIP5 and CMIP6 projections. Earth Syst. Dyn., 13, 321–340.

Danecek, P., Auton, A., Abecasis, G., Albers, C.A., Banks, E., DePristo, M.A., et al. (2011). The variant call format and VCFtools. Bioinformatics, 27, 2156–2158.

Danecek, P., Bonfield, J.K., Liddle, J., Marshall, J., Ohan, V., Pollard, M.O., et al. (2021). Twelve years of SAMtools and BCFtools. GigaScience, 10, giab008.

Davis, R.M. (1967). Parasitism by newly-transformed anadromous sea lampreys on landlocked salmon and other fishes in a coastal Maine Lake. Trans. Am. Fish. Soc., 96, 11–16.

Delaval, A., Frost, M., Bendall, V., Hetherington, S.J., Stirling, D., Hoarau, G., et al. (2022). Population and seascape genomics of a critically endangered benthic elasmobranch, the blue skate *Dipturus batis*. Evol. Appl., 15, 78–94.

Deutsch, C., Penn, J.L. & Lucey, N. (2024). Climate, oxygen, and the future of marine biodiversity. Annu. Rev. Mar. Sci., 16, 217–245.

Deutsch, C., Penn, J.L. & Seibel, B. (2020). Metabolic trait diversity shapes marine biogeography. Nature, 585, 557–562.

Diffenbaugh, N.S., Pal, J.S., Giorgi, F. & Gao, X. (2007). Heat stress intensification in the Mediterranean climate change hotspot. Geophys. Res. Lett., 34, 2007GL030000.

Docker, M.F. & Potter, I.C. (2019). Life history evolution in lampreys: Alternative migratory and feeding types. In: Lampreys: Biology, conservation and control (ed. Docker, M.F.). Springer Netherlands, Dordrecht, pp. 287–409.

Duarte, C.M., Hendriks, I.E., Moore, T.S., Olsen, Y.S., Steckbauer, A., Ramajo, L., et al. (2013). Is ocean acidification an open-ocean syndrome? Understanding anthropogenic impacts on seawater pH. Estuaries Coasts, 36, 221–236.

Ellis, N., Smith, S.J. & Pitcher, C.R. (2012). Gradient forests: Calculating importance gradients on physical predictors. Ecology, 93, 156–168.

Ferreira-Martins, D., Wilson, J.M., Kelly, S.P., Kolosov, D. & McCormick, S.D. (2021). A review of osmoregulation in lamprey. J. Gt. Lakes Res., 47, S59–S71.

Ford, M. (2024). *Petromyzon marinus*. IUCN Red List Threat. Species 2024 ET16781A58298056.

Frichot, E. & François, O. (2015). LEA: An R package for landscape and ecological association studies. Methods Ecol. Evol., 6, 925–929.

Gaggiotti, O.E., Bekkevold, D., Jørgensen, H.B.H., Foll, M., Carvalho, G.R., Andre, C., et al. (2009). Disentangling the effects of evolutionary, demographic, and environmental factors influencing genetic structure of natural populations: Atlantic herring as a case study. Evolution, 63, 2939–2951.

Gagnaire, P.-A., Normandeau, E., Côté, C., Møller Hansen, M. & Bernatchez, L. (2012). The genetic consequences of spatially varying selection in the panmictic American eel (*Anguilla rostrata*). Genetics, 190, 725–736.

Gilmour, K.M. & Perry, S.F. (2009). Carbonic anhydrase and acid–base regulation in fish. J. Exp. Biol., 212, 1647–1661.

Guo, B., Li, Z. & Merilä, J. (2016). Population genomic evidence for adaptive differentiation in the Baltic Sea herring. Mol. Ecol., 25, 2833–2852.

Habibzadeh, N., Ghoddousi, A., Bleyhl, B. & Kuemmerle, T. (2021). Rear-edge populations are important for understanding climate change risk and adaptation potential of threatened species. Conserv. Sci. Pract., 3, e375.

Haenel, Q., Roesti, M., Moser, D., MacColl, A.D.C. & Berner, D. (2019). Predictable genome-wide sorting of standing genetic variation during parallel adaptation to basic versus acidic environments in stickleback fish. Evol. Lett., 3, 28–42.

Hardisty, M.W. (1986). *Petromyzon marinus* Linnaeus, 1758. In: Holčík, J. (ed) The freshwater fishes of Europe, vol 1, part 1, Petromyzontiformes. Aula, Wiesbaden, pp 92–116.

Hardisty, M.W. & Potter, I.C. (1971). The general biology of adult lampreys. In: The biology of lampreys (eds. Hardisty M.W. & Potter I.C.). Academic Press, New York, pp. 127–206.

Hare, J.A., Morrison, W.E., Nelson, M.W., Stachura, M.M., Teeters, E.J., Griffis, R.B., et al. (2016). A vulnerability assessment of fish and invertebrates to climate change on the Northeast U.S. continental shelf. PLOS ONE, 11, e0146756.

Hay, A.C., Sandoval-Castillo, J., Cooke, G.M., Chao, N.L. & Beheregaray, L.B. (2022). Riverscape genomics clarifies neutral and adaptive evolution in an Amazonian characin fish (T*riportheus albus*). Front. Ecol. Evol., 10, 825406.

Hendriks, I.E., Duarte, C.M., Olsen, Y.S., Steckbauer, A., Ramajo, L., Moore, T.S., et al. (2015). Biological mechanisms supporting adaptation to ocean acidification in coastal ecosystems. Estuar. Coast. Shelf Sci., 152, A1–A8.

Hess, J.E., Campbell, N.R., Close, D.A., Docker, M.F. & Narum, S.R. (2013). Population genomics of Pacific lamprey: Adaptive variation in a highly dispersive species. Mol. Ecol., 22, 2898–2916.

Hess, J.E., Caudill, C.C., Keefer, M.L., McIlraith, B.J., Moser, M.L. & Narum, S.R. (2014). Genes predict long distance migration and large body size in a migratory fish, Pacific lamprey. Evol. Appl., 7, 1192–1208.

Hess, J.E., Delomas, T.A., Jackson, A.D., Kosinski, M.J., Moser, M.L., Porter, L.L., et al. (2022). Pacific lamprey translocations to the Snake River boost abundance of all life stages. Trans. Am. Fish. Soc., 151, 263–296.

Hess, J.E., Lampman, R.T., Jackson, A.D., Sween, T., Jim, L., McClain, N., et al. (2023). The return of the adult Pacific Lamprey offspring from translocations to the Columbia River. North Am. J. Fish. Manag., 43, 1531–1552.

Hess, J.E., Smith, J.J., Timoshevskaya, N., Baker, C., Caudill, C.C., Graves, D., et al. (2020). Genomic islands of divergence infer a phenotypic landscape in Pacific lamprey. Mol. Ecol., 29, 3841–3856.

Hume, J.B., Bennis, S., Bruning, T., Docker, M.F., Good, S., Lampman, R., et al. (2024). Evaluation of larval sea lamprey Petromyzon marinus growth in the laboratory: Influence of temperature and diet. Aquac. Res., 2024, 1–11.

Jiang, L.-Q., Carter, B.R., Feely, R.A., Lauvset, S.K. & Olsen, A. (2019). Surface ocean pH and buffer capacity: Past, present and future. Sci. Rep., 9, 18624.

Johnson, M.T.J. & Munshi-South, J. (2017). Evolution of life in urban environments. Science, 358, eaam8327.

Kawecki, T.J. & Ebert, D. (2004). Conceptual issues in local adaptation. Ecol. Lett., 7, 1225–1241.

King, J.J. & O' Gorman, N. (2018). Initial observations on feeding juvenile sea lamprey (*Petromyzon marinus* L.) in Irish lakes. Biol. Environ. Proc. R. Ir. Acad., 118B, 113.

Kjesbu, O.S., Sundby, S., Sandø, A.B., Alix, M., Hjøllo, S.S., Tiedemann, M., et al. (2022). Highly mixed impacts of near-future climate change on stock productivity proxies in the North East Atlantic. Fish Fish., 23, 601–615.

Lampman, R.T., Maine, A.N., Moser, M.L., Arakawa, H. & Neave, F.B. (2021). Lamprey aquaculture successes and failures: A path to production for control and conservation. J. Gt. Lakes Res., 47, S201–S215.

Láruson, Á.J., Fitzpatrick, M.C., Keller, S.R., Haller, B.C. & Lotterhos, K.E. (2022). Seeing the forest for the trees: Assessing genetic offset predictions from gradient forest. Evol. Appl., 15, 403–416.

Lassalle, G., Béguer, M., Beaulaton, L. & Rochard, E. (2008). Diadromous fish conservation plans need to consider global warming issues: An approach using biogeographical models. Biol. Conserv., 141, 1105–1118.

Lassalle, G. & Rochard, E. (2009). Impact of twenty-first century climate change on diadromous fish spread over Europe, North Africa and the Middle East. Glob. Change Biol., 15, 1072–1089.

Lejeusne, C., Chevaldonné, P., Pergent-Martini, C., Boudouresque, C.F. & Pérez, T. (2010). Climate change effects on a miniature ocean: the highly diverse, highly impacted Mediterranean Sea. Trends Ecol. Evol., 25, 250–260.

Li, H. & Durbin, R. (2009). Fast and accurate short read alignment with Burrows–Wheeler transform. Bioinformatics, 25, 1754–1760.

Limburg, K.E. & Waldman, J.R. (2009). Dramatic declines in North Atlantic diadromous fishes. BioScience, 59, 955–965.

Maar, M., Møller, E.F., Larsen, J., Madsen, K.S., Wan, Z., She, J., et al. (2011). Ecosystem modelling across a salinity gradient from the North Sea to the Baltic Sea. Ecol. Model., 222, 1696–1711.

Maitland, P.S. (1980). Review of the ecology of lampreys in Northern Europe. Can. J. Fish. Aquat. Sci., 37, 1944–1952.

Maitland, P.S., Morris, K.H. & East, K. (1994). The ecology of lampreys (Petromyzonidae) in the Loch Lomond area. Hydrobiologia, 290, 105–120.

Maitland, P.S., Renaud, C.B., Quintella, B.R., Close, D.A. & Docker, M.F. (2015). Conservation of native lampreys. In: Lampreys: Biology, conservation and control (ed. Docker, M.F.). Springer Netherlands, Dordrecht, pp. 375–428.

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., et al. (2010). The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res., 20, 1297–1303.

McLaren, W., Gil, L., Hunt, S.E., Riat, H.S., Ritchie, G.R.S., Thormann, A., et al. (2016). The Ensembl Variant Effect Predictor. Genome Biol., 17, 122.

Melzner, F., Gutowska, M.A., Langenbuch, M., Dupont, S., Lucassen, M., Thorndyke, M.C., et al. (2009). Physiological basis for high CO_2; tolerance in marine ectothermic animals: Pre-adaptation through lifestyle and ontogeny? Biogeosciences, 6, 2313–2331.

Moser, M.L., Hume, J.B., Aronsuu, K.K., Lampman, R.T. & Jackson, A.D. (2019). Lamprey reproduction and early life history: Insights from artificial propagation. In: Lampreys: Biology, conservation and control (ed. Docker, M.F.). Springer Netherlands, Dordrecht, pp. 187–245.

NatureServe (2024). NatureServe network biodiversity location data accessed through NatureServe explorer [web application]. http://explorer.natureserve.org

Nikinmaa, M. (1993). Haemoglobin function in intact *Lampetra fluviatilis* erythrocytes. Respir. Physiol., 91, 283–293.

Nikinmaa, M. & Mattsoff, L. (1992). Effects of oxygen saturation on the CO_2 transport properties of *Lampetra* red cells. Respir. Physiol., 87, 219–230.

Novak, B.J., Phelan, R. & Weber, M. (2021). U.S. conservation translocations: Over a century of intended consequences. Conserv. Sci. Pract., 3, e394.

Oksanen, J., Simpson, G., Blanchet, F., Kindt, R., Legendre, P., Minchin, P., et al. (2022). vegan: Community ecology package. R package version 2.6-4. https://CRAN.R-project.org/package=vegan.

Palumbi, S.R. (2001). Humans as the world's greatest evolutionary force. Science, 293, 1786–1790.

Pereira, A.M., Jonsson, B., Johannsson, M., Robalo, J.I. & Almada, V.C. (2012). Icelandic lampreys (*Petromyzon marinus*): where do they come from? Ichthyol. Res., 59, 83–85.

Poloczanska, E.S., Brown, C.J., Sydeman, W.J., Kiessling, W., Schoeman, D.S., Moore, P.J., et al. (2013). Global imprint of climate change on marine life. Nat. Clim. Change, 3, 919–925.

Pörtner, H.O. & Knust, R. (2007). Climate change affects marine fishes through the oxygen limitation of thermal tolerance. Science, 315, 95–97.

Potter, I.C., Gill, H.S., Renaud, C.B. & Haoucher, D. (2015). The taxonomy, phylogeny, and distribution of lampreys. In: Lampreys: Biology, conservation and control (ed. Docker, M.F.). Springer Netherlands, Dordrecht, pp. 35–73.

Privé, F., Luu, K., Vilhjálmsson, B.J. & Blum, M.G.B. (2020). Performing highly efficient genome scans for local adaptation with R package pcadapt version 4. Mol. Biol. Evol., 37, 2153–2154.

Raven, J., Caldeira, K., Elderfield, H., Hoegh-Guldberg, O., Liss, P., Riebesell, U., et al. (2005). Ocean acidification due to increasing atmospheric carbon dioxide. The Royal Society, Policy document 12/05.

Rehm, E.M., Olivas, P., Stroud, J. & Feeley, K.J. (2015). Losing your edge: Climate change and the conservation value of range-edge populations. Ecol. Evol., 5, 4315–4326.

Renaud, C.B. & Cochran, P.A. (2019). Post-metamorphic feeding in lampreys. In: Lampreys: Biology, conservation and control (ed. Docker, M.F.). Springer Netherlands, Dordrecht, pp. 247–285.

Riccioni, G., Stagioni, M., Landi, M., Ferrara, G., Barbujani, G. & Tinti, F. (2013). Genetic structure of bluefin tuna in the Mediterranean Sea correlates with environmental variables. PLoS ONE, 8, e80105.

Sala, E. & Knowlton, N. (2006). Global marine biodiversity trends. Annu. Rev. Environ. Resour., 31, 93–122.

Screen, J.A. (2014). Arctic amplification decreases temperature variance in northern mid- to high-latitudes. Nat. Clim. Change, 4, 577–582.

Silva, S., Servia, M.J., Vieira-Lanero, R. & Cobo, F. (2013a). Downstream migration and hematophagous feeding of newly metamorphosed sea lampreys (*Petromyzon marinus* Linnaeus, 1758). Hydrobiologia, 700, 277–286.

Silva, S., Servia, M.J., Vieira-Lanero, R., Nachón, D.J. & Cobo, F. (2013b). Haematophagous feeding of newly metamorphosed European sea lampreys *Petromyzon marinus* on strictly freshwater species. J. Fish Biol., 82, 1739–1745.

Spice, E.K., Goodman, D.H., Reid, S.B. & Docker, M.F. (2012). Neither philopatric nor panmictic: Microsatellite and mtDNA evidence suggests lack of natal homing but limits to dispersal in Pacific lamprey. Mol. Ecol., 21, 2916–2930.

Steffen, W., Persson, Å., Deutsch, L., Zalasiewicz, J., Williams, M., Richardson, K., et al. (2011). The Anthropocene: From global change to planetary stewardship. AMBIO, 40, 739–761.

Supek, F., Bošnjak, M., Škunca, N. & Šmuc, T. (2011). REVIGO summarizes and visualizes long lists of gene ontology terms. PLoS ONE, 6, e21800.

Timoshevskaya, N., Eşkut, K.İ., Timoshevskiy, V.A., Robb, S.M.C., Holt, C., Hess, J.E., et al. (2023). An improved germline genome assembly for the sea lamprey *Petromyzon marinus* illuminates the evolution of germline-specific chromosomes. Cell Rep., 42, 112263.

Törönen, P., Medlar, A. & Holm, L. (2018). PANNZER2: A rapid functional annotation web server. Nucleic Acids Res., 46, W84–W88.

Tufts, B.L., Bagatto, B. & Cameron, B. (1992). In vivo analysis of gas transport in arterial and venous blood of the sea lamprey *Petromyzon marinus*. J. Exp. Biol., 169, 105–119.

Tyberghein, L., Verbruggen, H., Pauly, K., Troupin, C., Mineur, F. & De Clerck, O. (2012). Bio-ORACLE: A global environmental dataset for marine species distribution modelling. Glob. Ecol. Biogeogr., 21, 272–281.

United Nations Department of Economic and Social Affairs, Population Division (2022). World population prospects 2022: Summary of results. UN DESA/POP/2022/TR/NO. 3.

Van Leeuwen, T.E., Cote, D., Pretty, C., Townley, J., Poole, R., Dempson, B., et al. (2021). Incursions of sea lamprey, *Petromyzon marinus*, and striped bass, *Morone saxatilis,* in Labrador waters: Episodic events or evidence of a northward range expansion? J. Fish Biol., 99, 2066–2070.

Waldman, J.R. & Quinn, T.P. (2022). North American diadromous fishes: Drivers of decline and potential for recovery in the Anthropocene. Sci. Adv., 8, eabl5486.

Wang, C.J., Michael Hudson, J., Lassalle, G. & Whitesel, T.A. (2021). Impacts of a changing climate on native lamprey species: From physiology to ecosystem services. J. Gt. Lakes Res., 47, S186–S200.

Ward, D.L., Clemens, B.J., Clugston, D., Jackson, A.D., Moser, M.L., Peery, C., et al. (2012). Translocating adult Pacific lamprey within the Columbia River basin: State of the science. Fisheries, 37, 351–361.

WGS500 Consortium, Rimmer, A., Phan, H., Mathieson, I., Iqbal, Z., Twigg, S.R.F., et al. (2014). Integrating mapping-, assembly- and haplotype-based approaches for calling variants in clinical sequencing applications. Nat. Genet., 46, 912–918.

Wu, T., Hu, E., Xu, S., Chen, M., Guo, P., Dai, Z., et al. (2021). clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. The Innovation, 2, 100141.

Table 4.1. Anadromous sea lamprey samples and associated metadata.

Sampling Site	Location	Abbreviation	Coordinates	Samples	Collection Date	Life Stage	Preservation Method	Collaborator	Affiliation
Close to Vestra-Horn	Iceland	ICE_67 ICE_68	64.162500, -14.476333	2	2009	Adult	Tissue - ethanol	Benóný Jónsson, Magnús Jóhannsson	Hafrannsóknastofnun (MFRI)
Close to Vestman Islands	Iceland	ICE_69	63.347633, -20.087850	1	2009	Adult	Tissue - ethanol	Benóný Jónsson, Magnús Jóhannsson	Hafrannsóknastofnun (MFRI)
Ytri Ranga River	Iceland	ICE_66	63.818367, -20.418817	1	2009	Adult	Tissue - ethanol	Benóný Jónsson, Magnús Jóhannsson	Hafrannsóknastofnun (MFRI)
At sea Heradsfloi	Iceland	ICE_70 - ICE_74	65.657933, -14.135033	5	2012	Adult	Tissue - ethanol	Benóný Jónsson, Magnús Jóhannsson	Hafrannsóknastofnun (MFRI)
River Rhine	Germany	PMGE2020	48.83316932, 8.10988970295832	8	2020	Adult & larvae	Fin clips - ethanol	Jan Baer	Landwirtschaftliches Zentrum Baden-Württemberg (LAZBW)
Lough Conn – Moy*	Ireland	Ireland_01 - Ireland_09	N/A	9	2010, 2011, 2016, 2017	Adult	Fin clips - ethanol	Fiona Bracken	University College Dublin

River Mulkear - Shannon	Ireland	Ireland_11 - Ireland_27	N/A	10	2010, 2011	Adult	Fin clips - ethanol	Fiona Bracken	University College Dublin
Lough Derg - Shannon*	Ireland	Ireland_28 - Ireland_37	N/A	8	2010, 2011, 2017	Adult	Fin clips - ethanol	Fiona Bracken	University College Dublin
Oir River	France Atlantic	SEL	N/A	7	2015, 2017, 2018	Adult	Fin clips - ethanol	Guillaume Evanno	Institut national de recherche pour l'agriculture, l'alimentation et l'environnement (INRAE)
Scorff River	France Atlantic	SC	N/A	8	2018	Adult	Fin clips - ethanol	Guillaume Evanno	Institut national de recherche pour l'agriculture, l'alimentation et l'environnement (INRAE)
Loire River	France Atlantic	PEMA	N/A	6	2013	Adult	Fin clips - ethanol	Guillaume Evanno	Institut national de recherche pour l'agriculture, l'alimentation et l'environnement (INRAE)
Vouga River Basin	Portugal	VOU	N/A	10	2021	Adult	Fin clips - ethanol	Catarina Mateus	Universidade de Évora
Mondego River Basin	Portugal	MON	N/A	8	2020	Adult	Fin clips - ethanol	Catarina Mateus	Universidade de Évora
Tejo River Basin	Portugal	TEJO	N/A	9	2021	Adult	Fin clips - ethanol	Catarina Mateus	Universidade de Évora
Guadiana River Basin	Portugal	GUA	N/A	10	2020, 2021	Adult	Fin clips - ethanol	Catarina Mateus	Universidade de Évora
Hogvadsan	Sweden	SWE	N/A	7	2021	Adult	Fin clips - ethanol	Mikael Svensson, Elisabeth Thysell	Sveriges lantbruksuniversitet (SLU), Länsstyrelsen i Hallands län
Rhône Estuary	France Mediterranean	FRMED	N/A	6	2003, 2010, 2015	Adult & NA	Fin clips - ethanol	Guillaume Evanno	Institut national de recherche pour l'agriculture, l'alimentation et l'environnement (INRAE)
Connecticut River	Massachusetts	CTR	41.28477, -72.342589	12	2019	Adult	Fin clips – frozen carcasses	Ted Castro-Santos	U.S. Geological Survey (USGS)
Richibucto River	New Brunswick	RI	46.692742, -64.842909	12	2019	Larvae	Fin clips – frozen carcasses	Mike Wilkie	Wilfrid Laurier University
Deer Creek	Maryland	DR2021	39.6748228, -76.4511499	2	2021	Adult	Fin clips - ethanol	Thomas Evans	St. Mary’s College of Maryland

Falling Branch	Maryland	FB2021	39.6909291, -76.4265029	2	2021	Adult	Fin clips - ethanol	Thomas Evans	St. Mary's College of Maryland
Little Patuxent River	Maryland	PR2021	39.134949, -76.82557	4	2021	Larvae	Fin clips - ethanol	Thomas Evans	St. Mary's College of Maryland
Pee Dee River**	North Carolina	NCSM_69176 NCSM_69181	34.922800, -80.018530 34.856270, -79.916770	2	2011	NA	Tissue - ethanol	Gabriela M.Hogue, Michael Fisk	North Carolina Museum of Natural Sciences, NC Wildlife Resources Commission
Neuse River**	North Carolina	NCSM_46040	35.349, -78.1501	1	2007	NA	Tissue - ethanol	Gabriela M.Hogue, Michael Fisk	North Carolina Museum of Natural Sciences, NC Wildlife Resources Commission
Roanoke River**	North Carolina	NCSM_38060 NCSM_105133 NCSM_97531 NCSM_97532	36.328700, -77.578100 36.477120, -77.639720 36.481510 to 36.468700, -77.643820 to -77.636830	4	2018, 2019, 2024	Adult & NA	Tissue - ethanol	Gabriela M.Hogue, Michael Fisk	North Carolina Museum of Natural Sciences, NC Wildlife Resources Commission
Pamlico River, Tar River**	North Carolina	NCSM 53260	35.955400, -77.780700	1	2009	Adult	Tissue - ethanol	Gabriela M.Hogue, Michael Fisk	North Carolina Museum of Natural Sciences, NC Wildlife Resources Commission

* Locations at which feeding in freshwater prior to outmigration has been observed, ** Samples from museum collection (North Carolina Museum of Natural Sciences Ichthyology Collection.

Table 4.2. Table showing effective population size (Ne) estimates from Chapter 3 and maladaptation likelihood per location. For Ireland, Portugal and France, where maladaptation likelihood was estimated at higher resolution for distinct river basins, I interpret results at the country level so they are comparable to Ne estimates that are informative at the broader scale. The NA value observed for New Brunswick is considered an artifact of large population size.

Location	**Effective Population Size (Ne)**	**Maladaptation likelihood**
New Brunswick	NA	High
Sweden	~ 61	High
Ireland	~ 1186	Moderate
France Atlantic	~ 2260	Moderate

Germany	~ 3	Moderate
Portugal	~ 6997	Moderate
Iceland	~ 358	Moderate
Massachusetts	~ 9758	Moderate
North Carolina	~ 151	Low
Mediterranean (French coast)	~ 24	Low
Maryland	~ 38	Low

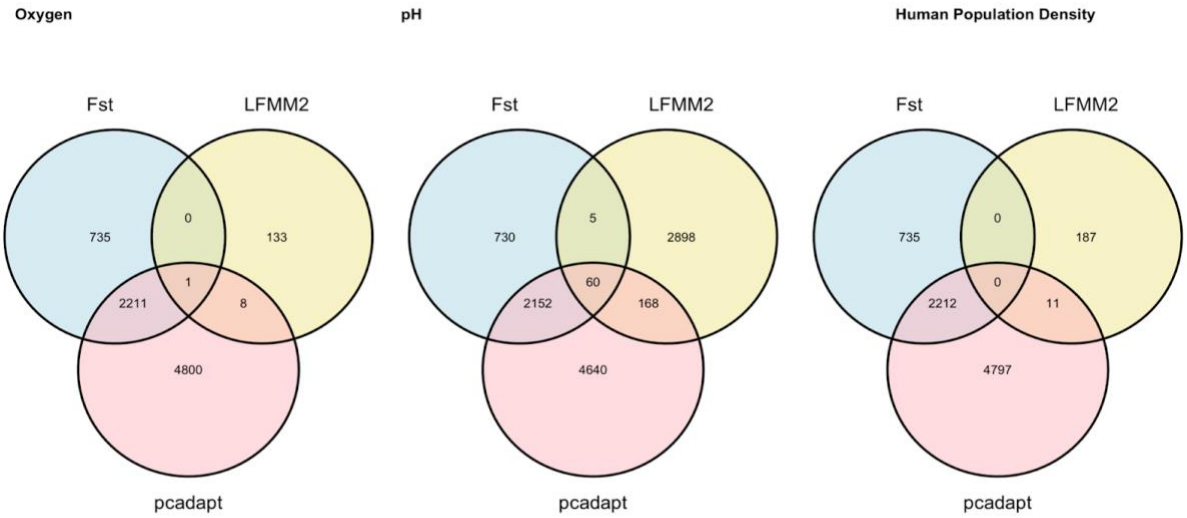

Figure 4.1. Outlier SNPs connected to surface dissolved oxygen, surface pH and human population density detected across all three methods (LFMM2, pcadapt and F_{ST} outliers).

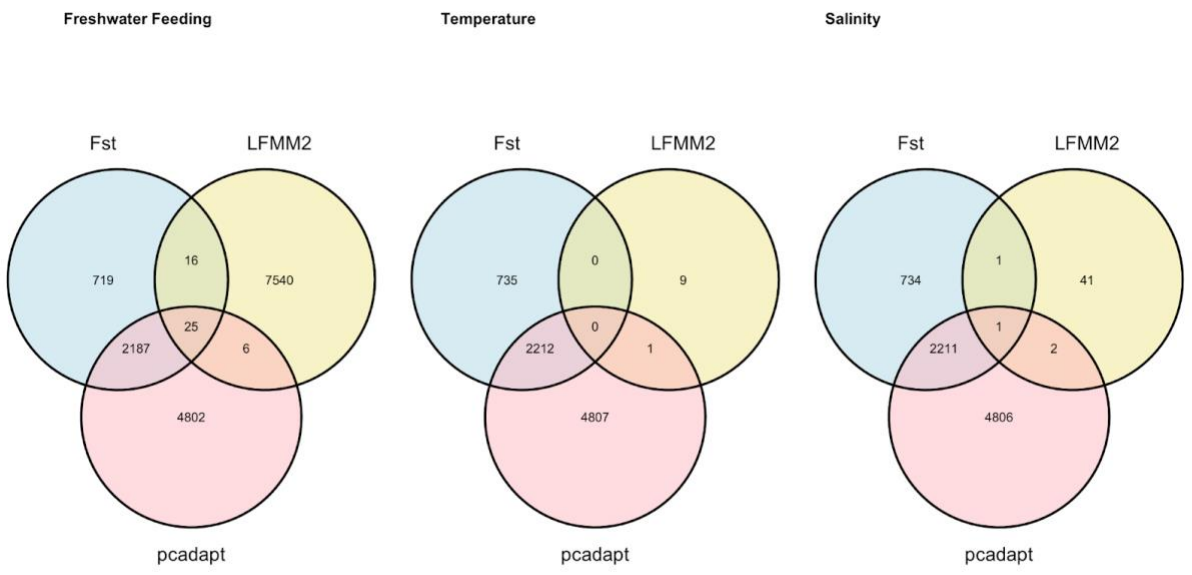

Figure 4.2. Outlier SNPs connected to freshwater feeding, sea surface temperature and surface salinity detected across all three methods (LFMM2, pcadapt and F_{ST} outliers).

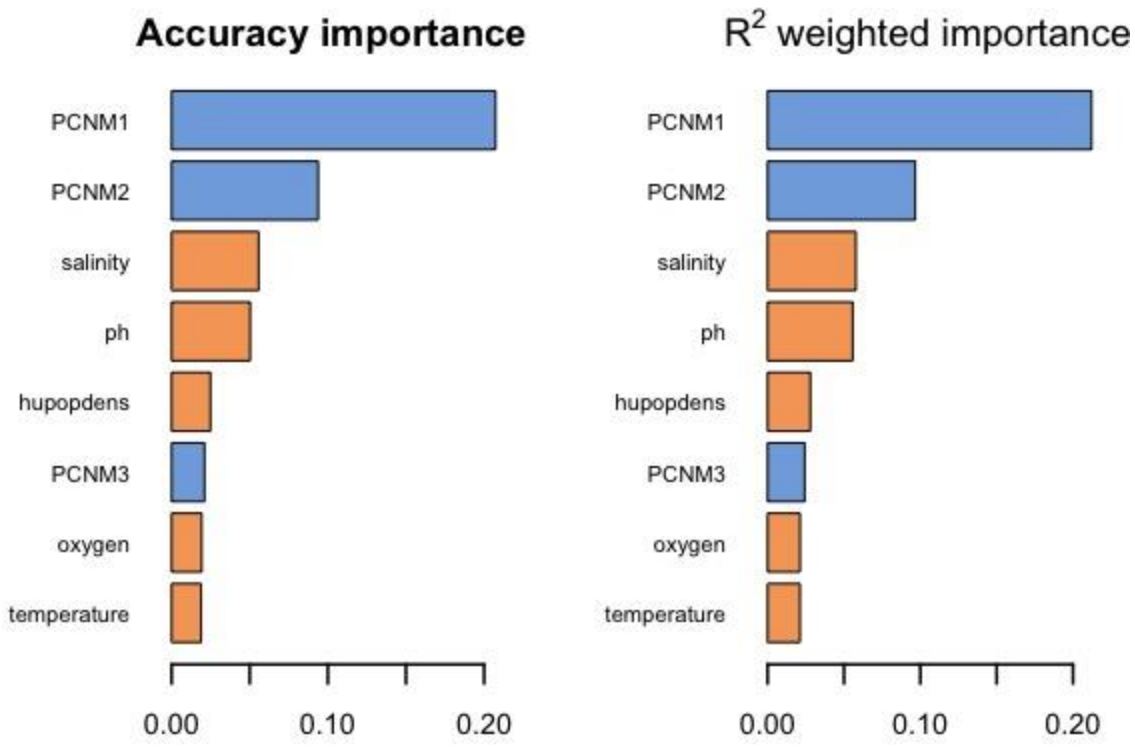

Figure 4.3. Fitted gradient forest model predictor variables and significant Principal Coordinates of Neighbourhood Matrix (PCNMs) ranked by importance in the model. Presented as mean accuracy importance and mean importance weighted by species R^2. Spatial predictor variables in blue and environmental predictor variables in orange.

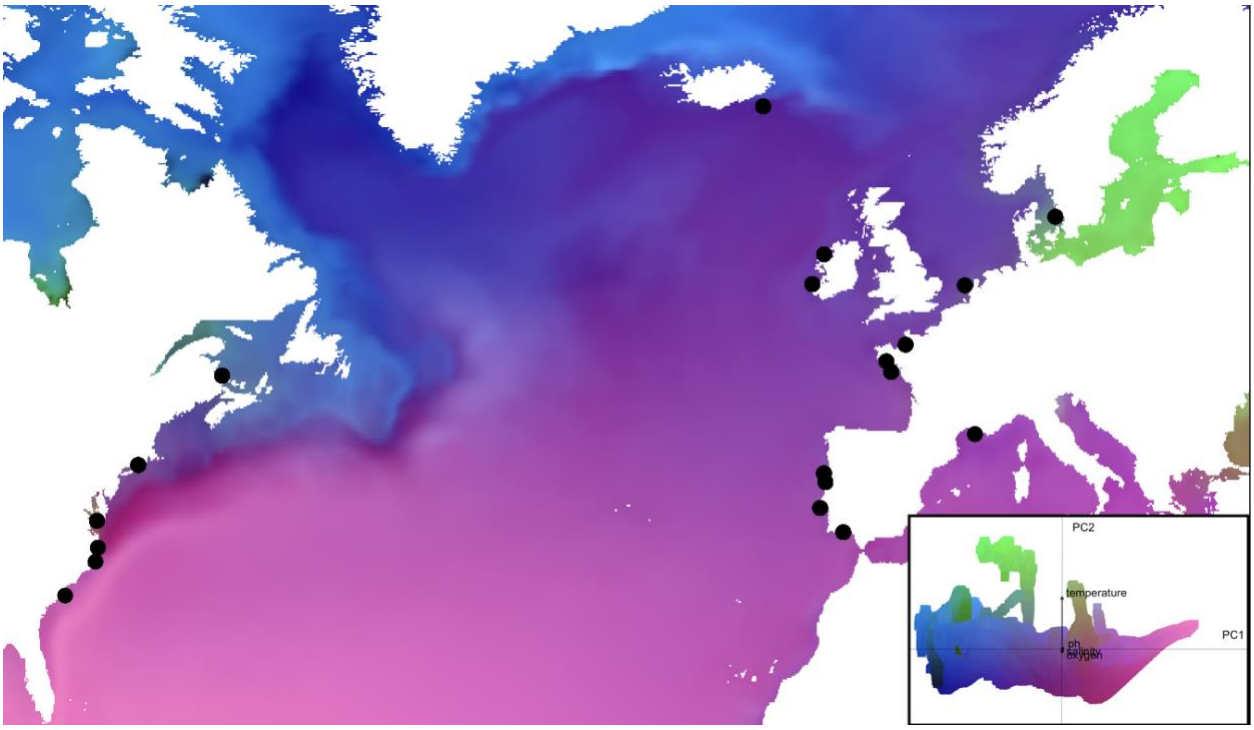

Figure 4.4. Map of allelic turnover across space and environmental gradients in salinity, pH, oxygen and temperature. Locations with similar colors are expected to be genetically similar based on the gradient forest model. The color scale is created along the principal component axes of the gradient forest model (bottom right).

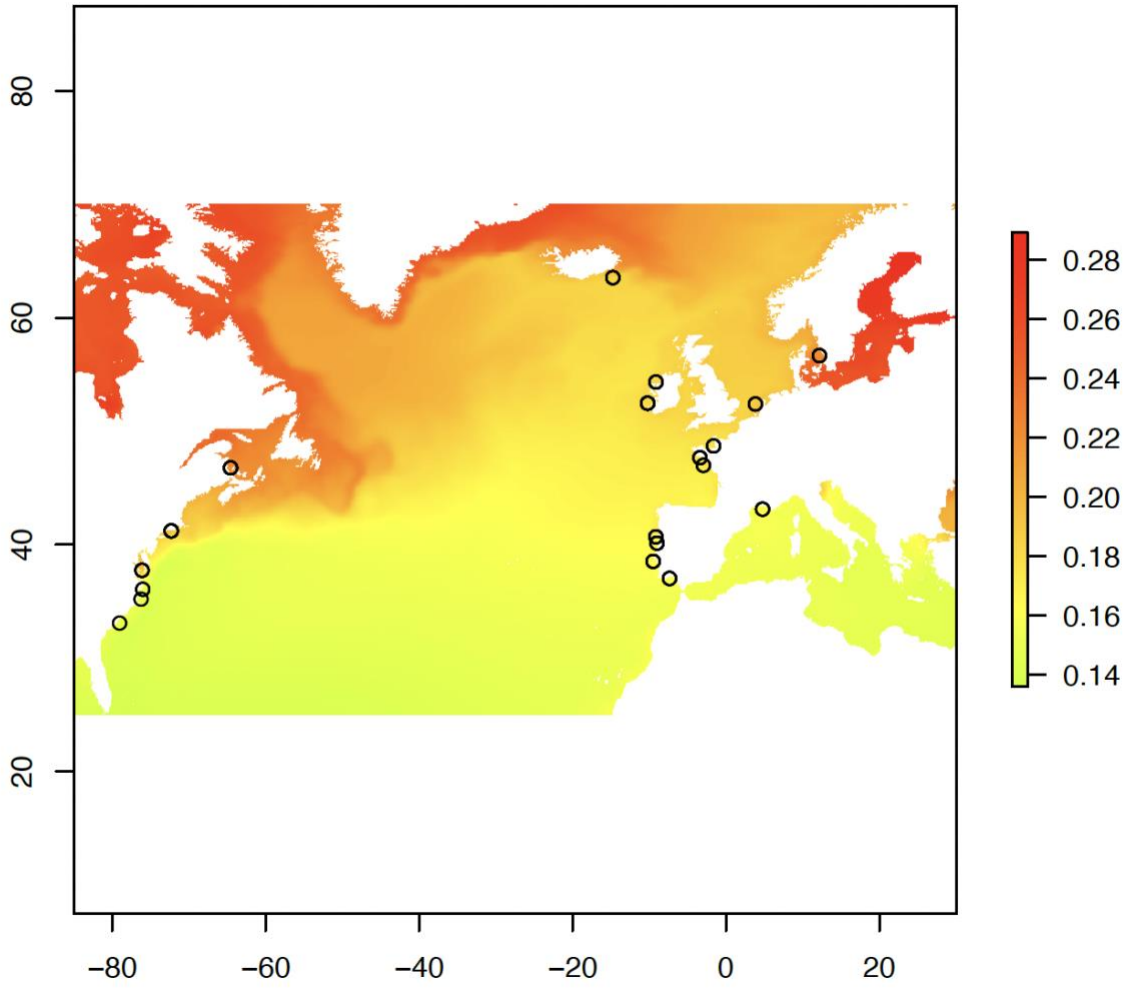

Figure 4.5. Map of genetic offsets across space and environmental gradients in salinity and temperature under a RCP6.0 scenario for the years 2090-2100. Higher values corresponding to darker colors represent higher genetic offsets while lower values corresponding to lighter colors represent lower genetic offsets. Higher genetic offsets represent higher vulnerability while lower genetic offsets represent lower vulnerability.

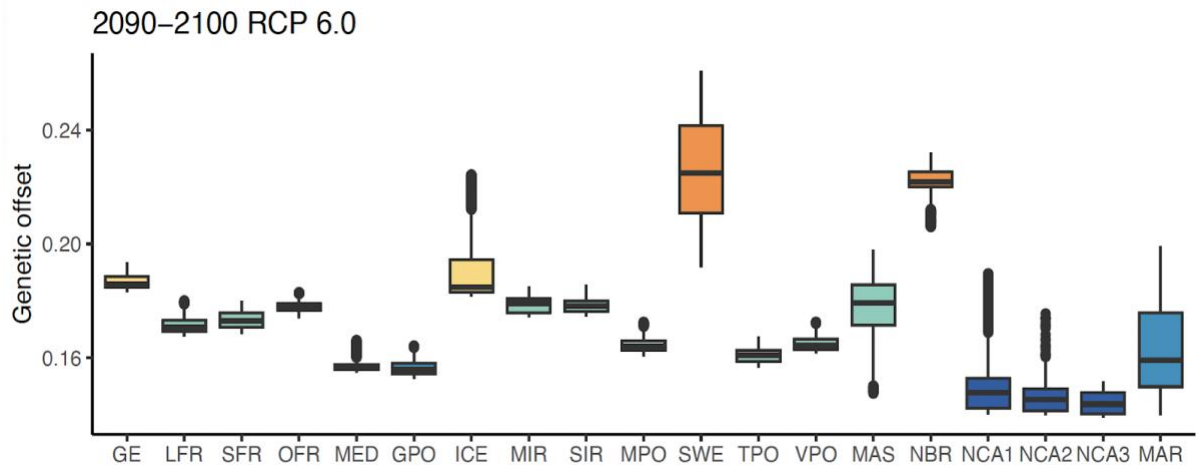

Figure 4.6. Boxplots of per site genetic offsets under a RCP6.0 scenario for the years 2090-2100. Sites represented as following on the x axis: Germany (GE), River Loire France (LFR), River Scorff France (SFR), River Oir France (OFR), France Mediterranean Coast (MED), Guadiana River Portugal (GPO), Iceland (ICE), River Moy Ireland (MIR), River Shannon (SIR), Mondego River Portugal (MPO), Sweden (SWE), Tejo River Portugal (TPO), Vouga River Portugal (VPO), Massachusetts (MAS), New Brunswick (NBR), Roanoke River North Carolina (NCA1), Pamlico/Tar & Neuse Rivers North Carolina & (NCA2), Pee Dee River North Carolina (NCA3), Maryland (MAR). Higher values on the y axis represent higher genetic offsets while lower values represent lower genetic offsets. Higher genetic offsets indicate higher vulnerability while lower genetic offsets indicate lower vulnerability.

CHAPTER 5: DISCUSSION

Macrogenetic studies can take advantage of the vast amount of data available from previous studies and repurpose it to push the boundaries of environmental science (Halpern *et al.* 2023; Leigh *et al*. 2021). The logistics of traditional sampling methods can be overcome and research made more accessible. However, my intended global study (Chapter 2) does not fully represent global genetic diversity but constitutes the best representation possible with currently available data. Gaps in my data are likely part of the sampling bias known to affect the study of biodiversity (Hughes *et al.* 2021; Titley *et al.* 2017). There is clear geographic bias in my data, with data availability likely linked to a country's economic resources. Differences in data archiving practices could have prevented some datasets from being programmatically detected, and not all data is publicly available. Polar regions and deep-sea environments are not well represented due to logistical difficulties in collecting these data. While economically important commercially fished species made up the majority of species used in my study (36), those supporting minor commercial fisheries (15) and those not commercially fished (22) were well represented (See Supplemental Information 1.2). When aiming to integrate genetic diversity into conservation policy, effort should be made for data availability in underrepresented regions and equity in biodiversity research. Additionally, making the most of existing data by encouraging and optimizing data reuse methods (Binley *et al.* 2024) and open data practices (Poisot *et al.* 2019; Renaut *et al.* 2018) can help fill in knowledge and data gaps and inform conservation (Binley *et al.* 2024; Roche *et al.* 2022).

At the species level, my thesis produces knowledge on anadromous native populations of sea lamprey, an often underappreciated species in some parts of its range. Due to invasive landlocked populations in the Laurentian Great Lakes being the subject of extensive control

programs (Marsden & Siefkes 2019), sea lamprey often get a bad reputation, especially in North America. Meanwhile in Portugal and elsewhere in Europe, sea lamprey are a highly valued, culturally important species (Almeida *et al.* 2021; Docker *et al.* 2015), showing the paradox in how we view and place value on nature. The importance of conservation of anadromous native sea lamprey and other native lamprey species has been extensively supported (Clemens & Wang 2021; Hume *et al.* 2021; Lucas *et al.* 2021; Maitland *et al.* 2015), and my thesis adds a population genomic approach to conservation across the entire sea lamprey native range. It is worth noting that this study only considers population structure, current and future adaptation in regards to the marine environment. As an anadromous species, sea lamprey spend a significant part of their life cycle in freshwater. Therefore, to fully understand and plan for sea lamprey conservation, it is encouraged to replicate this study with a focus on freshwater habitat.

There could be other evolutionary forces underlying global genetic diversity patterns. Testing these hypotheses is unfortunately outside the scope of this study but provides exciting new prospects for the field of macrogenetics. It is yet to be tested how global biogeographic regions (Spalding *et al.* 2007) can explain genetic diversity patterns in marine fish. Meanwhile in terrestrial mammals, populations closer to biogeographic boundaries are less genetically diverse, more highly differentiated and with lower effective population sizes indicating limited adaptive potential in biogeographic transitional zones (Schmidt *et. al* 2022). Understanding the relationship between regional biotic and abiotic conditions and evolutionary processes could inform protected area planning and biodiversity conservation. Environmental heterogeneity is also considered an important driver of evolution (Stein *et al.* 2014) and found to influence genetic variation within marine fish and invertebrate species at the local level, in the form of temperature range (Dalongeville *et al.* 2022). Meanwhile, stable environments are thought to

harbor higher genetic diversity (Hewitt 2000), and environmental stability has been connected to higher within-species genetic diversity in South African marine invertebrate species (Nielsen *et al.* 2021). The connection between genetic diversity and stability is important to consider under future climate change related disruptions. It would also be beneficial to model genetic diversity under future projected climatic conditions to investigate how relationships could change globally. It could also be valuable for future studies to take into account the seasonality of environmental conditions. Life history traits could also underlie global patterns and could be tested for; vertical distribution, migration type and body length are found to drive genetic diversity in Mediterranean fish species (Dalongeville *et al.* 2016).

Associations found between global gene diversity, allelic richness and candidate environmental and human disturbance metrics can be interpreted in regard to their conservation relevance across historical time scales (Caballero *et al.* 2010). Allelic richness, affected by effective population size and more susceptible to bottlenecks, is a better indicator of contemporary evolution (Luikart *et al.* 1998) while expected heterozygosity is less sensitive to their influence and better reflects past evolutionary events (Nei *et al.* 1975). As primary productivity and land area are correlated with both gene diversity and allelic richness, these natural processes seem to influence past and contemporary evolution. Temperature, dissolved oxygen and species richness seem to be more strongly influenced by recent evolutionary processes. The influence of urbanization on marine environments is relevant across all time scales, indicating the extent of human interference on nearby ecosystems and potential difficulty for contemporary evolutionary processes to compensate. Declines associated with cumulative human impacts are more visible at a recent time scale. As human intervention amplifies across global oceans contemporary evolutionary mechanisms could struggle to compensate for genetic

diversity declines. The negative association between genetic diversity and fishing activity is relevant at a more distant time scale. Contemporary evolutionary processes may have a buffering effect, encouraging hopes for adaptive management as a mitigation strategy as the dietary needs of a growing human population continue to increase. Further research on the precise mechanisms involved and their intensity will be needed to test these theories.

My thesis has further proven the conservation utility of genetic data at both the species and global level, and I discuss implications for consideration in protected areas. The value of including a conservation genetic approach in marine protected area design is increasingly recognized despite lack of large-scale implementation (Laikre *et al.* 2016; Sandström *et al.* 2016; Von der Heyden 2009). Lack of uniform data coverage prevents inclusion of genetic diversity metrics in protected areas design globally, but opportunities arise at the local level (Paz-Vinas *et al.* 2023). Well-designed marine protected areas can further conservation goals by preserving biodiversity within and contributing to maintaining biodiversity outside their borders. These conservation goals can be achieved by ensuring adequate connectivity within and between designated protected areas, and connectivity between protected and non-protected areas (Xuereb *et al.* 2019). Studies have shown that accounting for genetic connectivity (Cros *et al.* 2017; van der Meer *et al.* 2015) and defining genetically distinct management units (Coelho *et al.* 2024) in marine protected area design and placement can benefit genetic diversity, and subsequently biodiversity conservation. Genetic diversity within and outside of reserves can be maintained by gene flow (Huserbråten *et al.* 2013; Pujolar *et al.* 2013) and can assist population recovery (Aalto *et al.* 2019; Munguía-Vega *et al.* 2015). Maintaining high standing genetic diversity is essential to consider in marine protected areas, as vulnerability to future threats increases with lower adaptive potential (Taboada *et al.* 2023). Planning protected areas to include already

existing adaptive genetic variation will ensure populations are better prepared against future environmental changes (Von der Heyden 2017).

Species level information on anadromous sea lamprey could be considered in new or existing protected areas throughout their native range. My study found that sea lamprey in the North Atlantic constitute two genetically distinct admixed populations, one on each Atlantic coast, further confirming previous studies (Almada *et al.* 2008; Bryan *et al.* 2005; Genner *et al* 2012; Rodriguez-Munõz *et al.* 2004). Sea lamprey in the Mediterranean are genetically distinct from their East Atlantic counterparts but retain high levels of gene flow. Therefore, each Atlantic coast and the Mediterranean will need consideration of its unique circumstances and gene flow between the East Atlantic and Mediterranean should be maintained. An anadromous species that disperses widely, sea lamprey could benefit from considering both their marine and freshwater habitat when planning for protected areas. This could help protect sea lamprey across all life stages. However, the main threats to anadromous sea lamprey, barriers to migration and habitat degradation (Maitland *et al.* 2015), are predominant in their freshwater habitat. While it is worth considering marine areas at locations where sea lamprey are most threatened when planning reserves, protecting freshwater spawning and larval habitats, migration routes and their connectivity is essential to maintaining healthy populations. Sea lamprey freshwater habitat and migration routes at locations found to be vulnerable in the present day such as Sweden, Maryland, Germany and the Mediterranean Sea (see Chapter 3), and locations vulnerable under future environmental conditions such as Sweden and New Brunswick (see Chapter 4) could be recommended for protection and restoration.

Protected areas located in open waters may not have a direct effect on sea lamprey but could be indirectly beneficial by protecting their host species, future changes in host abundance

already considered a threat to parasitic lampreys under climate change (Maitland *et al.* 2019). Sea lamprey survival is thought to depend on host density (Hansen *et al.* 2016), but since parasitic lampreys tend to not be host-specific (Renaud & Cochran 2019) they are more likely vulnerable to changes in host abundance rather than host distribution (Maitland *et al.* 2019). Protected areas located across the Atlantic to Mediterranean transition zone could help maintain high rates of gene flow observed between these distinct populations. This gene flow is important to be maintained as it could be contributing to maintaining genetic diversity and therefore adaptive potential in the Mediterranean, a historically small sea lamprey population that is currently threatened (see Chapter 3). Connectivity through gene flow will help maintain standing genetic diversity, with high standing genetic diversity connected to adaptive potential (Barrett & Schluter 2008). As sea lamprey within each Atlantic coast are admixed, gene flow from in-reserve sea lamprey would benefit sea lamprey inhabiting neighboring locations outside reserves. High connectivity within each Atlantic coast provides the opportunity to benefit across its entire native range through strategic placement of a protected area network catering to the needs of diadromous fish species. Information produced on anadromous sea lamprey adds to literature available for North Atlantic diadromous species (e.g., King et al. 2007; Lehnert et al. 2020; Nikolic et al. 2020; Nilsson et al. 2001; Rougemont & Bernatchez 2018). Literature on population structure and connectivity could be used to inform future protected area design and planning in regards to diadromous fishes as a whole. Diadromous species have complex, highly migratory life cycles that can make conservation multifaceted. Marine protected areas are limited in how much protection they can offer diadromous fish species that spend a significant amount of their life in fresh water. Therefore, it is valuable to consider designating specific marine to

freshwater protected areas for diadromous species, in areas where many species overlap, and introducing protocols specific to diadromous species during the reserve designing process.

While adaptive potential is typically connected to high standing genetic diversity (Barrett & Schluter 2008), genetic offsets are able to identify locations at which individuals are better or worse adapted to future environmental conditions (Láruson *et al.* 2022). Despite limitations to predictions of maladaptation, such as not accounting for future species range expansions (Rellstab *et al.* 2021) and the unpredictability of novel environments under future climate change (Lotterhos *et al.* 2021), genetic offsets can offer some insight into what the future holds. As protected areas should be designed to include current and maximize future adaptive potential (Von der Heyden 2017, Xuereb *et al.* 2019) and high levels of gene flow can facilitate the spread of these beneficial alleles (Whiteley *et al.* 2015), it is worth considering planning new or extending existing reserves to include locations where a species is forecasted to be better adapted. In the case of anadromous sea lamprey, more southern locations such as North Carolina, Maryland, Portugal and the Mediterranean are predicted to be better adapted to future environmental conditions and could be prioritized for inclusion into protected areas. Despite range contractions expected in the southern parts of their European range (Lassalle *et al.* 2008), and by extrapolation potentially their entire southern range, sea lamprey at these locations may already be better adapted to climate change due to selective pressures already underway. Locations expected to persevere such as the Minho River basin in Portugal (Lassalle *et al.* 2008) could be especially resilient in the future. Freshwater habitat loss and degradation is a known threat to sea lamprey in Portugal (Almeida *et al.* 2002), further encouraging protection of these critical habitats. While studies have not been undertaken for sea lamprey specifically in North Carolina and Maryland, habitat loss and degradation are known to threaten diadromous species

based on per state habitat assessments (Maryland Department of Natural Resources 2005; North Carolina Wildlife Resources Commission (NC WRC) 2005). In the Mediterranean diadromous species are vulnerable to habitat fragmentation and loss due to damming (Duarte *et al.* 2021)

My thesis has also identified genetic diversity and gene flow patterns for marine fish species globally. A first step in the practical use of these results for marine protected area planning and design would be to increase data coverage, as there are many information gaps especially in developing countries. Including additional taxa would be beneficial as accounting across taxonomic levels will enable ecosystem wide conservation. I uncovered that across global oceans, areas of high genetic diversity also display high gene flow. This finding indicates that higher exchange of alleles through gene flow could be linked to higher rates of evolution leading to higher genetic diversity. Thus, adaptive potential and connectivity are intrinsically connected in global marine fishes, genetic diversity mediated by gene flow underlying adaptive potential (Barrett & Schluter 2008). This finding highlights the importance of designing interconnected protected area networks (Gaines *et al.* 2010; Kininmonth *et al.* 2011). However, my study focused on global genetic diversity patterns, while genetic diversity at the local level remains unaccounted for. In order to realistically think about the practical application of genetic diversity metrics to protected area planning, per ocean and more localized genetic diversity patterns will also need to be accounted for.

My thesis has detected global patterns in genetic diversity across marine fishes, and uncovered their underlying environmental drivers and implications of human activity. Conservation implications and future research directions have been discussed, highlighting the amount of knowledge the emerging field of macrogenetics has to offer. I have undertaken a conservation genomic approach to account for anadromous sea lamprey population structure and

status, while uncovering the species' demographic and evolutionary trajectory in the deep and more recent past. Predictions regarding future vulnerability and ability to adapt have been made. Given the relationship between genetic diversity, connectivity, adaptive potential and species future resilience, this thesis has attempted to take advantage of both genomic and publicly archived data to identify future conservation applications at both the global and species level, hoping to assist efforts to safeguard the future of global oceans.

References

Aalto, E.A., Micheli, F., Boch, C.A., Espinoza Montes, J.A., Woodson, C.B. & De Leo, G.A. (2019). Catastrophic mortality, Allee effects, and marine protected areas. Am. Nat., 193, 391–408.

Almada, V.C., Pereira, A.M., Robalo, J.I., Fonseca, J.P., Levy, A., Maia, C., et al. (2008). Mitochondrial DNA fails to reveal genetic structure in sea-lampreys along European shores. Mol. Phylogenet. Evol., 46, 391–396.

Almeida, P., Quintella, B., Dias, N. & Andrade, N. (2002). The anadromous sea lamprey in Portugal: Biology and conservation perspectives. In: The biology of lampreys, symposium proceedings (eds. Moser, M., Bayer, J. & MacKinlay, D.). Presented at the international congress on the biology of fish, American Fisheries Society, 21–25 July, Vancouver, British Columbia, pp. 49–58.

Almeida, P.R., Arakawa, H., Aronsuu, K., Baker, C., Blair, S.-R., Beaulaton, L., et al. (2021). Lamprey fisheries: History, trends and management. J. Gt. Lakes Res., 47, S159–S185.

Barrett, R. & Schluter, D. (2008). Adaptation from standing genetic variation. Trends Ecol. Evol., 23, 38–44.

Binley, A.D., Vincent, J.G., Rytwinski, T., Soroye, P. & Bennett, J.R. (2024). Making the most of existing data in conservation research. Perspect. Ecol. Conserv., 22, 122–128.

Bryan, M.B., Zalinski, D., Filcek, K.B., Libants, S., Li, W. & Scribner, K.T. (2005). Patterns of invasion and colonization of the sea lamprey (*Petromyzon marinus*) in North America as revealed by microsatellite genotypes. Mol. Ecol., 14, 3757–3773.

Caballero, A., Rodríguez-Ramilo, S.T., Ávila, V. & Fernández, J. (2010). Management of genetic diversity of subdivided populations in conservation programmes. Conserv. Genet., 11, 409–419.

Clemens, B.J. & Wang, C.J. (2021). Dispelling misperceptions of native lampreys (*Entosphenus* and *Lampetra* spp.) in the Pacific northwest (USA). Conserv. Sci. Pract., 3, e402.

Coelho, J.F.R., Mendes, L.D.F., Di Dario, F., Carvalho, P.H., Dias, R.M., Lima, S.M.Q., et al. (2024). Integration of genomic and ecological methods inform management of an undescribed, yet highly exploited, sardine species. Proc. R. Soc. B Biol. Sci., 291, 20232746.

Cros, A., Toonen, R.J., Donahue, M.J. & Karl, S.A. (2017). Connecting Palau's marine protected areas: A population genetic approach to conservation. Coral Reefs, 36, 735–748.

Dalongeville, A., Andrello, M., Mouillot, D., Albouy, C. & Manel, S. (2016). Ecological traits shape genetic diversity patterns across the Mediterranean Sea: A quantitative review on fishes. J. Biogeogr., 43, 845–857.

Dalongeville, A., Nielsen, E.S., Teske, P.R. & Von Der Heyden, S. (2022). Comparative phylogeography in a marine biodiversity hotspot provides novel insights into evolutionary processes across the Atlantic-Indian Ocean transition. Divers. Distrib., 28, 2622–2636.

Docker, M.F., Hume, J.B. & Clemens, B.J. (2015). Introduction: A surfeit of lampreys. In: Lampreys: Biology, conservation and control (ed. Docker, M.F.). Springer Netherlands, Dordrecht, pp. 1–34.

Duarte, G., Segurado, P., Haidvogl, G., Pont, D., Ferreira, M.T. & Branco, P. (2021). Damn those damn dams: Fluvial longitudinal connectivity impairment for European diadromous fish throughout the 20th century. Sci. Total Environ., 761, 143293.

Gaines, S.D., White, C., Carr, M.H. & Palumbi, S.R. (2010). Designing marine reserve networks for both conservation and fisheries management. Proc. Natl. Acad. Sci., 107, 18286–18293.

Genner, M.J., Hillman, R., McHugh, M., Hawkins, S.J. & Lucas, M.C. (2012). Contrasting demographic histories of European and North American sea lamprey (*Petromyzon marinus*) populations inferred from mitochondrial DNA sequence variation. Mar. Freshw. Res., 63, 827.

Halpern, B.S., Boettiger, C., Dietze, M.C., Gephart, J.A., Gonzalez, P., Grimm, N.B., et al. (2023). Priorities for synthesis research in ecology and environmental science. Ecosphere, 14, e4342.

Hansen, M.J., Madenjian, C.P., Slade, J.W., Steeves, T.B., Almeida, P.R. & Quintella, B.R. (2016). Population ecology of the sea lamprey (*Petromyzon marinus*) as an invasive species in the Laurentian Great Lakes and an imperiled species in Europe. Rev. Fish Biol. Fish., 26, 509–535.

Hewitt, G. (2000). The genetic legacy of the Quaternary ice ages. Nature, 405, 907–913.

Hughes, A.C., Orr, M.C., Ma, K., Costello, M.J., Waller, J., Provoost, P., et al. (2021). Sampling biases shape our view of the natural world. Ecography, 44, 1259–1269.

Hume, J.B., Almeida, P.R., Buckley, C.M., Criger, L.A., Madenjian, C.P., Robinson, K.F., et al. (2021). Managing native and non-native sea lamprey (*Petromyzon marinus*) through anthropogenic change: A prospective assessment of key threats and uncertainties. J. Gt. Lakes Res., 47, S704–S722.

Huserbråten, M.B.O., Moland, E., Knutsen, H., Olsen, E.M., André, C. & Stenseth, N.Chr. (2013). Conservation, spillover and gene flow within a network of northern European marine protected areas. PLoS ONE, 8, e73388.

King, T.L., Verspoor, E., Spidle, A.P., Gross, R., Phillips, R.B., Koljonen, M., et al. (2007). Biodiversity and population structure. In: The Atlantic salmon (eds. Verspoor, E., Stradmeyer, L. & Nielsen, J.). Wiley, pp. 117–166.

Kininmonth, S., Beger, M., Bode, M., Peterson, E., Adams, V.M., Dorfman, D., et al. (2011). Dispersal connectivity and reserve selection for marine conservation. Ecol. Model., 222, 1272–1282.

Laikre, L., Lundmark, C., Jansson, E., Wennerström, L., Edman, M. & Sandström, A. (2016). Lack of recognition of genetic biodiversity: International policy and its implementation in Baltic Sea marine protected areas. Ambio, 45, 661–680.

Láruson, Á.J., Fitzpatrick, M.C., Keller, S.R., Haller, B.C. & Lotterhos, K.E. (2022). Seeing the forest for the trees: Assessing genetic offset predictions from gradient forest. Evol. Appl., 15, 403–416.

Lassalle, G., Béguer, M., Beaulaton, L. & Rochard, E. (2008). Diadromous fish conservation plans need to consider global warming issues: An approach using biogeographical models. Biol. Conserv., 141, 1105–1118.

Lehnert, S.J., Kess, T., Bentzen, P., Clément, M. & Bradbury, I.R. (2020). Divergent and linked selection shape patterns of genomic differentiation between European and North American Atlantic salmon (*Salmo salar*). Mol. Ecol., 29, 2160–2175.

Leigh, D.M., Van Rees, C.B., Millette, K.L., Breed, M.F., Schmidt, C., Bertola, L.D., et al. (2021). Opportunities and challenges of macrogenetic studies. Nat. Rev. Genet., 22, 791–807.

Lotterhos, K.E., Láruson, Á.J. & Jiang, L.-Q. (2021). Novel and disappearing climates in the global surface ocean from 1800 to 2100. Sci. Rep., 11, 15535.

Lucas, M.C., Hume, J.B., Almeida, P.R., Aronsuu, K., Habit, E., Silva, S., et al. (2021). Emerging conservation initiatives for lampreys: Research challenges and opportunities. J. Gt. Lakes Res., 47, S690–S703.

Luikart, G., Sherwin, W.B., Steele, B.M. & Allendorf, F.W. (1998). Usefulness of molecular markers for detecting population bottlenecks via monitoring genetic change. Mol. Ecol., 7, 963–974.

Maitland, P.S., Renaud, C.B., Quintella, B.R., Close, D.A. & Docker, M.F. (2015). Conservation of native lampreys. In: Lampreys: Biology, conservation and control (ed. Docker, M.F.). Springer Netherlands, Dordrecht, pp. 375–428.

Marsden, J.E. & Siefkes, M.J. (2019). Control of invasive sea lamprey in the Great Lakes, Lake Champlain, and Finger Lakes of New York. In: Lampreys: Biology, conservation and control (ed. Docker, M.F.). Springer Netherlands, Dordrecht, pp. 411–479.

Maryland Department of Natural Resources (2005). Maryland Wildlife Diversity Conservation Plan. Annapolis, Maryland. https://dnr.maryland.gov/wildlife/documents/swap/2005_mwdc_plan.pdf

Munguía-Vega, A., Sáenz-Arroyo, A., Greenley, A.P., Espinoza-Montes, J.A., Palumbi, S.R., Rossetto, M., et al. (2015). Marine reserves help preserve genetic diversity after impacts derived from climate variability: Lessons from the pink abalone in Baja California. Glob. Ecol. Conserv., 4, 264–276.

Nei, M., Maruyama, T. & Chakraborty, R. (1975). The bottleneck effect and genetic variability in populations. Evolution, 29, 1.

Nielsen, E.S., Beger, M., Henriques, R. & Von Der Heyden, S. (2021). Neither historical climate nor contemporary range fully explain the extant patterns of molecular diversity in marine species. J. Biogeogr., 48, 2629–2644.

Nikolic, N., Liu, S., Jacobsen, M.W., Jónsson, B., Bernatchez, L., Gagnaire, P., et al. (2020). Speciation history of European (*Anguilla anguilla*) and American eel (*A. rostrata*), analysed using genomic data. Mol. Ecol., 29, 565–577.

Nilsson, J., Gross, R., Asplund, T., Dove, O., Jansson, H., Kelloniemi, J., et al. (2001). Matrilinear phylogeography of Atlantic salmon (*Salmo salar* L.) in Europe and postglacial colonization of the Baltic Sea area. Mol. Ecol., 10, 89–102.

North Carolina Wildlife Resources Commission (NC WRC). (2005). North Carolina wildlife action plan.

Paz-Vinas, I., Vandergast, A., Schmidt, C., Leigh, D., Blanchet, S., Clark, R., et al. (2023). Uneven genetic data limits biodiversity assessments in protected areas globally.

Poisot, T., Bruneau, A., Gonzalez, A., Gravel, D. & Peres-Neto, P. (2019). Ecological data should not be so hard to find and reuse. Trends Ecol. Evol., 34, 494–496.

Pujolar, J.M., Schiavina, M., Di Franco, A., Melià, P., Guidetti, P., Gatto, M., et al. (2013). Understanding the effectiveness of marine protected areas using genetic connectivity patterns and Lagrangian simulations. Divers. Distrib., 19, 1531–1542.

Rellstab, C., Dauphin, B. & Exposito-Alonso, M. (2021). Prospects and limitations of genomic offset in conservation management. Evol. Appl., 14, 1202–1212.

Renaud, C.B. & Cochran, P.A. (2019). Post-metamorphic feeding in lampreys. In: Lampreys: Biology, conservation and control (ed. Docker, M.F.). Springer Netherlands, Dordrecht, pp. 247–285.

Renaut, S., Budden, A.E., Gravel, D., Poisot, T. & Peres-Neto, P. (2018). Management, archiving, and sharing for biologists and the role of research institutions in the technology-oriented age. BioScience, 68, 400–411.

Roche, D.G., O'Dea, R.E., Kerr, K.A., Rytwinski, T., Schuster, R., Nguyen, V.M., et al. (2022). Closing the knowledge-action gap in conservation with open science. Conserv. Biol., 36, e13835.

Rodríguez-Muñoz, R., Waldman, J.R., Grunwald, C., Roy, N.K. & Wirgin, I. (2004). Absence of shared mitochondrial DNA haplotypes between sea lamprey from North American and Spanish rivers. J. Fish Biol., 64, 783–787.

Rougemont, Q. & Bernatchez, L. (2018). The demographic history of Atlantic salmon (*Salmo salar*) across its distribution range reconstructed from approximate Bayesian computations: Atlantic salmon history and linked selection. Evolution, 72, 1261–1277.

Sandström, A., Lundmark, C., Jansson, E., Edman, M. & Laikre, L. (2016). Assessment of management practices regarding genetic biodiversity in Baltic Sea marine protected areas. Biodivers. Conserv., 25, 1187–1205.

Schmidt, C., Muñoz, G., Lancaster, L.T., Lessard, J., Marske, K.A., Marshall, K.E., et al. (2022). Population demography maintains biogeographic boundaries. Ecol. Lett., 25, 1905–1913.

Spalding, M.D., Fox, H.E., Allen, G.R., Davidson, N., Ferdaña, Z.A., Finlayson, M., et al. (2007). Marine ecoregions of the world: A bioregionalization of coastal and shelf areas. BioScience, 57, 573–583.

Stein, A., Gerstner, K. & Kreft, H. (2014). Environmental heterogeneity as a universal driver of species richness across taxa, biomes and spatial scales. Ecol. Lett., 17, 866–880.

Taboada, S., Whiting, C., Wang, S., Ríos, P., Davies, A.J., Mienis, F., et al. (2023). Long distance dispersal and oceanographic fronts shape the connectivity of the keystone sponge *Phakellia ventilabrum* in the deep northeast Atlantic. Front. Mar. Sci., 10, 1177106.

Titley, M.A., Snaddon, J.L. & Turner, E.C. (2017). Scientific research on animal biodiversity is systematically biased towards vertebrates and temperate regions. PLOS ONE, 12, e0189577.

Van Der Meer, M.H., Berumen, M.L., Hobbs, J.-P.A. & Van Herwerden, L. (2015). Population connectivity and the effectiveness of marine protected areas to protect vulnerable, exploited and endemic coral reef fishes at an endemic hotspot. Coral Reefs, 34, 393–402.

Von Der Heyden, S. (2009). Why do we need to integrate population genetics into South African marine protected area planning? Afr. J. Mar. Sci., 31, 263–269.

Von Der Heyden, S. (2017). Making evolutionary history count: Biodiversity planning for coral reef fishes and the conservation of evolutionary processes. Coral Reefs, 36, 183–194.

Whiteley, A.R., Fitzpatrick, S.W., Funk, W.C. & Tallmon, D.A. (2015). Genetic rescue to the rescue. Trends Ecol. Evol., 30, 42–49.

Xuereb, A., D'Aloia, C.C., Daigle, R.M., Andrello, M., Dalongeville, A., Manel, S., et al. (2019). Marine conservation and marine protected areas. In: Population genomics: Marine organisms, population genomics (eds. Oleksiak, M.F. & Rajora, O.P.). Springer International Publishing, Cham, pp. 423–446.